



SCHOOL of  
GRADUATE STUDIES  
EAST TENNESSEE STATE UNIVERSITY

East Tennessee State University  
**Digital Commons @ East  
Tennessee State University**

---

Electronic Theses and Dissertations

Student Works

---

December 1978

# An Analysis of Three Selected Mental Maturity Measures in Predicting Academic Achievement of Students for Educational Decision Makers

Robert M. Davis

*East Tennessee State University*

Follow this and additional works at: <https://dc.etsu.edu/etd>



Part of the [Educational Administration and Supervision Commons](#)

---

## Recommended Citation

Davis, Robert M., "An Analysis of Three Selected Mental Maturity Measures in Predicting Academic Achievement of Students for Educational Decision Makers" (1978). *Electronic Theses and Dissertations*. Paper 2846. <https://dc.etsu.edu/etd/2846>

This Dissertation - Open Access is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact [digilib@etsu.edu](mailto:digilib@etsu.edu).

7907823

DAVIS, ROBERT MARSHALL, JR.

AN ANALYSIS OF THREE SELECTED MENTAL MATURITY  
MEASURES IN PREDICTING ACADEMIC ACHIEVEMENT  
OF STUDENTS FOR EDUCATIONAL DECISION MAKERS.

EAST TENNESSEE STATE UNIVERSITY, ED.D., 1978

University  
Microfilms  
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106

AN ANALYSIS OF THREE SELECTED MENTAL MATURITY MEASURES  
IN PREDICTING ACADEMIC ACHIEVEMENT OF STUDENTS  
FOR EDUCATIONAL DECISION MAKERS

---

A Dissertation  
Presented to  
the Doctoral Advisory Committee of the Graduate School  
East Tennessee State University

---

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Education

---

by  
Robert Marshall Davis, Jr.

December 1978

APPROVAL

This is to certify that the Advanced Graduate Committee of

ROBERT MARSHALL DAVIS, JR.

met on the

28th day of November, 19 78.

The committee read and examined his dissertation, supervised his defense of it in an oral examination, and decided to recommend that his study be submitted to the Graduate Council and the Dean of the School of Graduate Studies in partial fulfillment of the requirements for the degree Doctor of Education.

A. Keith Durkin

Chairman, Advanced Graduate Committee

J. Howard Bowers

Charles W. Benbelt

W. Floyd Haggen

Shemar Ronald

Elizabeth L. McMahon

Dean, School of Graduate Studies

Robert Marshall Davis, Jr., B. A., St. Andrews Presbyterian College,  
June 1963.  
M.Ed., University of Virginia, June 1968.  
D.A.G.S. (Diploma of Advanced Graduate  
Study), University of Virginia, June 1974.  
Ed.D., East Tennessee State University,  
December 1978.

AN ANALYSIS OF THREE SELECTED MENTAL MATURITY MEASURES  
IN PREDICTING ACADEMIC ACHIEVEMENT OF STUDENTS  
FOR EDUCATIONAL DECISION MAKERS

Problem. The problem was to determine by comparison which of the three selected mental maturity measures was the best predictor of achievement in mathematics and reading as measured by the Peabody Individual Achievement Test.

Method. This was a descriptive correlational analysis study.

This study's sample was limited to no more than the first 200 students in grades one through five, and no more than the first 100 students in grades six through twelve referred to Special Services by their teachers, guidance counselor or principal.

The criteria used in drawing the sample from the population was that all of the potential 300 students referred were presumed to have some type of handicapped condition.

The standardized instruments used in this study included three mental maturity tests and one achievement test. The three predictors in the study were the Stanford-Binet Intelligence Scale (SBIS), the Wechsler Intelligence Scale for Children-Revised (WISC-R), and the Slosson Intelligence Test (SIT). The criterion used in the study was the Peabody Individual Achievement Test (PIAT).

The entire sample was tested using the SIT, SBIS, WISC-R, and PIAT by professionally qualified examiners and certified psychologists. The three mental maturity tests were administered within a three week period from the time of referral. The time schedule limits for the administration of the three selected mental maturity measures was from October 1977 through March 1978. The PIAT was administered during May of 1978. The level of significance established for this study was .05. The statistical test applied was the Pearson Product-Moment Correlation. The resulting  $\gamma$  statistic was the validity coefficient  $\gamma_{xy}$  which was the correlation between predictor and criterion. The higher the validity coefficient, the greater was the correlation between the two variables.

Findings and Conclusions. The study sample was composed of two groups. The first group represented 140 elementary students in grades one through five. The second group represented 91 secondary students in grades six through twelve. The total sample population was 231 students. It was found in Groups I and II that the learning disability handicapped category had the highest referral percentage showing it was 80 percent of the sample for Group I and 57 percent of the sample for Group II, giving an overall 70 percent of the entire sample referred.

The primary hypothesis stated there was no significant difference in the predictive ability between the SIT, SBIS, and the WISC-R when compared to academic achievement in mathematics and reading as measured on the PIAT.

The sub-hypotheses stated there were no significant differences between mental ages, as derived from SIT, SBIS, and the WISC-R, and the MA as derived from grade placement on mathematics and reading of the PIAT.

Both null hypothesis were rejected.

Comparisons of the groups with IQ achievement raw scores with those groups with IQ achievement mental age scores revealed that the WISC-R 75 percent of the time and the SBIS 25 percent of the time were the best predictors and held the most substantial relationship for IQ achievement raw scores and IQ achievement mental age scores.

The WISC-R provided for a more global and gestalt appraisal of those verbal and non-verbal types of tasks that were necessary in the processing of information relative to reception, association, and expression which was imperative in the acquisition of mathematics and reading skills.

In addition there was an educationally significant discrepancy between their estimated intellectual potential and actual level of performance, and there existed even a greater discrepancy between their actual level of performance and their expected grade placement level. For the most part these specific learning disabilities were related to basic disorders in the learning process which the WISC-R so well detects as part of the differential diagnostic process.

. . . . .

Dissertation prepared under the direction of Dr. A. Keith Turkett, Dr. J. Howard Bowers, Dr. Charles W. Burkett, Dr. W. Lloyd Graunke, and Dr. Thomas G. Ronald.

Dedication

to

my loved ones

Jane

Cathy Karen Bruce

who taught me

more than they

will ever know

to

my beloved parents

Jack and Rosa Lee

who taught me

the meaning of

determination and perseverance

to

Dr. A. Keith Turkett, Chairman

whose example as a professional and Christian

taught me the true meaning of

a work fit to live for;

a self fit to live with;

a faith fit to live by.

## ACKNOWLEDGEMENTS

The pursuit and completion of this study would not have been possible without the assistance and cooperation of many persons.

The writer wishes to express his immeasurable appreciation to Dr. A. Keith Turkett, adviser and chairman of the advisory committee, Dr. J. Howard Bowers, Dr. Charles W. Burkett, Dr. W. Lloyd Graunke, and Dr. Thomas G. Ronald for their competence, devotion, keen insight, thoughtful and constructive criticism, and friendly counsel.

The writer is grateful to Dr. Clyde Orr and Dr. William Caskey who took time to give valuable reactions and suggestions during the course of this study.

The writer wishes to express his gratitude to Dr. Elizabeth L. McMahan for her extraordinary insight into the study as well as her advice to me during the pursuit of the doctoral program.

The writer wishes to express thanks to Martha Littleford for her expertise in typing, to Amelia Schumaier for her assistance in the computer analysis, and to Martha Hauff for her expertise in editing and journalistic implementation.

Above all, I am deeply indebted to my wife Jane, daughters Cathy and Karen, and son Bruce who made tremendous sacrifices while I pursued the doctorate. If it had not been for their tremendous love, understanding, support, patience, and tolerance, the writer could not have pursued or attained this degree. Finally, the writer is indebted to his mother and father whose love and sustained support was a source of inspiration.



## CONTENTS

	Page
APPROVAL . . . . .	ii
ABSTRACT . . . . .	iii
DEDICATION . . . . .	v
ACKNOWLEDGEMENTS . . . . .	vi
LIST OF TABLES . . . . .	xi
 Chapter	
1. INTRODUCTION . . . . .	1
The Problem . . . . .	3
Statement of the Problem . . . . .	3
Definitions of Terms . . . . .	3
Criterion-Related Validity . . . . .	3
Peabody Individual Achievement Test (PIAT) . . . . .	4
Predictive Validity . . . . .	4
Public Law 94-142 (P.L. 94-142) . . . . .	4
Slosson Intelligence Test (SIT) . . . . .	4
Stanford-Binet Intelligence Scale (SBIS) . . . . .	5
Wechsler Intelligence Scale for Children-Revised (WISC-R) . . . . .	5
Emotionally Disturbed . . . . .	5
Handicapped Children . . . . .	6
Hearing Impaired . . . . .	6
Learning Disabled . . . . .	6
Mentally Retarded . . . . .	7

Chapter	Page
Multiple Handicapped . . . . .	7
Physically Handicapped . . . . .	7
Special Education . . . . .	7
Speech Impaired . . . . .	8
Visually Impaired . . . . .	8
Limitations . . . . .	8
Significance of the Study . . . . .	9
Background . . . . .	9
Justification . . . . .	10
Sources of Data . . . . .	13
Assumptions . . . . .	13
Hypotheses . . . . .	13
A. Primary Hypothesis . . . . .	13
B. Sub-Hypotheses . . . . .	14
Procedures . . . . .	14
Related Studies . . . . .	14
Treatment of the Data . . . . .	18
Summary . . . . .	19
2. REVIEW OF RELATED LITERATURE . . . . .	21
Introduction . . . . .	21
Administrative Decision-Making . . . . .	23
Nature of Intelligence and Psychological Testing . . . . .	26
Measurement of Intelligence . . . . .	33
Intelligence Tests . . . . .	35
Stanford-Binet Intelligence Scale . . . . .	35
Validity and Reliability . . . . .	39

Chapter	Page
Classification System . . . . .	40
Wechsler Intelligence Scale for Children and the Wechsler Intelligence Scale for Children-Revised . .	42
Validity and Reliability . . . . .	45
Classification System . . . . .	46
Revision of the WISC . . . . .	48
Slosson Intelligence Test . . . . .	49
Validity and Reliability . . . . .	56
Classification System . . . . .	56
Standardized Achievement Tests . . . . .	57
Peabody Individual Achievement Test . . . . .	60
Validity and Reliability . . . . .	62
Classification System . . . . .	62
Significant Studies . . . . .	63
Summary . . . . .	75
3. METHODOLOGY . . . . .	76
Introduction . . . . .	76
Population . . . . .	76
The Sample . . . . .	77
Tests Identified and Described . . . . .	78
Stanford-Binet Intelligence Scale (SBIS) . . . . .	79
Wechsler Intelligence Scale for Children- Revised (WISC-R). . . . .	79
Slosson Intelligence Test (SIT) . . . . .	81
Peabody Individual Achievement Test (PIAT) . . . . .	81
Method Applied . . . . .	82
Summary . . . . .	83

Chapter	Page
4. ANALYSIS OF DATA . . . . .	84
Introduction . . . . .	84
The Sample . . . . .	84
Primary Hypothesis . . . . .	89
Sub-hypothesis . . . . .	91
Summary . . . . .	95
5. SUMMARY . . . . .	96
Introduction . . . . .	96
Findings . . . . .	97
Conclusions . . . . .	101
Recommendations . . . . .	107
BIBLIOGRAPHY . . . . .	109
APPENDIXES	
A. Procedures for the Identification, Evaluation, Confirmation, and Placement of Special Education Students . . . . .	115
B. Procedures for the Reevaluation of Special Education Students . . . . .	119
C. Collected Data Worksheet . . . . .	123
VITA . . . . .	125

## LIST OF TABLES

Table	Page
1. Distribution of the 1937 Standardization Group . . . . .	39
2. Distribution of the 1949 Standardization Group . . . . .	45
3. Changes in the WISC-R . . . . .	50
4. Relation of WISC-R IQ's Scales Scores and Scaled Scores Classification to Deviations from the Mean and Percentile Ranks . . . . .	53
5. Percentile Ranks for IQ's . . . . .	54
6. IQ Classification Chart . . . . .	55
7. Percentages of the Sample Population Handicapped Categories . . . . .	85
8. Numbers and Percentages by Schools of the Sample Population as to Time of Referral . . . . .	86
9. Mean Age (Criterion) for Both Groups, Raw Score Means and Standard Deviations of the Predictors for Both Groups, and Adjusted Mental Age Means and Standard Deviations of the Predictors and Criterion for Both Groups . . . . .	87
10. Pearson Product-Moment Correlation Coefficients Between PIAT Raw Scores on the Mathematics and Reading Subtests and the SIT, SBIS, and WISC-R(V), (P), (FS) IQ Raw Scores . . . . .	90
11. Pearson Product-Moment Correlation Coefficients Between PIAT Mental Age Scores on the Mathematics and Reading Subtests and the SIT, SBIS, and WISC-R(V), (P), and (FS) Mental Age Scores . . . . .	92
12. Pearson Product-Moment Correlation Coefficients for Groups I and II Among the Intelligence Tests Raw Scores (N = 231) . . . . .	94

## Chapter 1

### INTRODUCTION

Stephen J. Knezevich stated that school administration was

. . . a social process concerned with identifying, maintaining, stimulating, controlling, and unifying formally and informally organized human and material energies within an integrated system designed to accomplish predetermined objectives. . . .<sup>1</sup>

. . . An educational institution that requires a pattern of administration to propel it efficiently and effectively toward realization of its goals, to maintain and sustain it on an even keel, to steer it through often uncharted problem areas, and to keep it energized and prepared to weather challenges of fast-changing times.<sup>2</sup>

Daniel Griffiths suggested that the function of administration was to develop and regulate the decision-making process because decision-making was generally recognized as the heart of the administrative process.<sup>3</sup> The writings of Chester Barnard, Herbert A. Simon, and Robert T. Livingston stressed the processes necessary to put the decision into operation and implementation.<sup>4</sup> Harriet Talmage suggested that educators continually engage in activities concerning evaluation because the results

---

<sup>1</sup>Stephen J. Knezevich, Administration of Public Education (3d ed.; New York: Harper and Row, 1975), p. 12.

<sup>2</sup>Knezevich, p. 9.

<sup>3</sup>Daniel E. Griffiths, Administrative Theory (Englewood Cliffs, New Jersey: Prentice-Hall, 1959), pp. 73-75.

<sup>4</sup>Chester I. Barnard, Functions of the Executive (Cambridge, Massachusetts: Harvard University Press, 1938), p. 215; Herbert A. Simon, Administrative Behavior (New York: Macmillan, 1950), p. 1; Robert Teviot Livingston, "The Theory of Organization and Management," Transactions of the ASME, May 1953, p. 659.

of evaluation lead to decision making.<sup>5</sup> William Roe and Thelbert Drake stressed the relationship between the evaluation process and decision making.<sup>6</sup> Leon Lessinger was concerned with accountability as it related to educational goals and outcomes.<sup>7</sup> Knezevich stressed accountability as being relevant to educational goals as well as providing a clarification of the relationship between inputs and outputs.<sup>8</sup>

Differences among these viewpoints of administration suggested that the decision-making process needed to be more accurate, that the implementation of evaluation strategies was a process of prediction, and that accountability was related to the educational goals and outcomes achieved through its programs and activities.

All decisions involved prediction.<sup>9</sup> Prediction was adequately achieved by using tests which were valuable and contribute to the decision-making process.<sup>10</sup> The concern about individuals was to do something about them, individually or collectively.<sup>11</sup>

A predictor is the variable used to predict future performance on the basis of a given personal characteristic,

---

<sup>5</sup>Harriet Talmage, Statistics as a Tool for Educational Practitioners (Berkeley, California: McCutchan, 1976), pp. v-ix.

<sup>6</sup>William H. Roe and Thelbert L. Drake, The Principalship (New York: Macmillan, 1974), pp. 165-166.

<sup>7</sup>Leon M. Lessinger, "Accountability: Present Forces and Future Concerns," New Directions for Education, I (Spring, 1973), 1-9.

<sup>8</sup>Knezevich, p. 599.

<sup>9</sup>Lee J. Cronbach, Essentials of Psychological Testing (3d ed.; New York: Harper and Row, 1970), p. 22.

<sup>10</sup>Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), p. 23.

<sup>11</sup>Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education (3d ed.; New York: John Wiley and Sons, 1969), p. 8.

ability, or other measurable trait . . . a criterion is the variable used as the standard for measuring a performance after a period of instruction or treatment . . . criterion-related validity is based on the relationship between a predictor and a criterion . . . permits inferences to be made about one variable (the criterion) from scores obtained on another variable (the predictor).<sup>12</sup>

The type of criterion-related validity used in this study was predictive validity. "Predictive validity is concerned with inferring future performance from present performance."<sup>13</sup>

Implicit in the concept of criterion-related validity is the idea that tests are used as part of a decision-making process . . . its usefulness is an index of its relative contribution over and above that of other measures and sources of information to increased decision-making accuracy.<sup>14</sup>

### The Problem

#### Statement of the Problem

The problem was to determine by comparison which of the three selected mental maturity measures was the best predictor of achievement in mathematics and reading as measured by the Peabody Individual Achievement Test.

### Definitions of Terms

#### Criterion-Related Validity

Criterion-related validity "permits inferences to be made about one variable (the criterion) from scores obtained on another variable (the predictor)."<sup>15</sup>

---

<sup>12</sup>Talmage, pp. 75, 112-113.

<sup>13</sup>Talmage, p. 113.

<sup>14</sup>Frederick Brown, Principles of Educational and Psychological Testing (Hinsdale, Illinois: The Dryden Press, 1970), pp. 103-104.

<sup>15</sup>Talmage, p. 113.



Peabody Individual Achievement  
Test (PIAT)

This test was a wide-range individual screening test of achievement with age ranges from kindergarten through adult. The test surveyed an individual's level of educational attainment in the areas of mathematics, reading recognition, reading comprehension, spelling, and general information.<sup>16</sup>

Predictive Validity

Predictive validity "is predicated on the extent to which an instrument or entering performance will predict future success on a criterion measure inferring future performance from present performance."<sup>17</sup>

Public Law 94-142 (P.L. 94-142)

The essence of Public Law 94-142, The Education for All Handicapped Children Act, was that "after September 1, 1978 . . . it will be a violation of federal law . . . to deny any handicapped child a free, appropriate public education and a variety of accompanying rights."<sup>18</sup>

Slosson Intelligence Test (SIT)

This test was a short individual test of intelligence. Many of the items were adapted from the Stanford-Binet, Form L-M, 1960, third

---

<sup>16</sup> Lloyd M. Dunn and Frederick C. Markwardt, Jr., Manual for the Peabody Individual Achievement Test (Circle Pines, Minnesota: American Guidance Service, 1970), pp. 18-31.

<sup>17</sup> Harriet Talmage, Statistics as a Tool for Educational Practitioners (Berkeley, California: McCutchan, 1976), pp. 113-115.

<sup>18</sup> Scottie Torres, ed., A Primer on Individualized Education Programs for Handicapped Children (Reston, Virginia: The Foundation for Exceptional Children, 1977), p. 1.

revision. It was used as an individual screening instrument with individuals from ages two weeks to adult.<sup>19</sup>

Stanford-Binet Intelligence  
Scale (SBIS)

This test was an individual test of intelligence. It was newly normed (1972). The third revision (1960) form L-M consisted of one form used to measure intelligence of individuals from age two to adult.<sup>20</sup>

Wechsler Intelligence Scale  
for Children-Revised (WISC-R)

This test was an individual test of intelligence. It measured verbal and performance tasks, and it was a revision (1974) and restandardization of the 1949 WISC. The age range was six years to sixteen years, eleven months.<sup>21</sup>

The writer recognized that the following definitions were peculiar to the State of Virginia.

Emotionally Disturbed

Children who are emotionally disturbed demonstrate one or more of the following characteristics to a marked extent and over a period of time:

(a) An inability to learn which cannot be explained by intellectual, sensory, or health factors.

---

<sup>19</sup> Richard L. Slosson, Slosson Intelligence Test for Children and Adults (New York: Slosson Educational Publications, 1963), pp. 16-17.

<sup>20</sup> Lewis M. Terman and Maud A. Merrill, Manual for the Stanford-Binet Intelligence Scale (3d ed.; Boston: Houghton Mifflin, 1973), pp. 5, 20, 353-61.

<sup>21</sup> David Wechsler, Manual for the Wechsler Intelligence Scale for Children-Revised (New York: Psychological Corporation, 1974), pp. 5-10.

(b) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers.

(c) Inappropriate types of behavior or feeling under normal conditions.

(d) A general, pervasive mood of unhappiness or depression.

(e) A tendency to develop physical symptoms, pains, or fears associated with personal or school problems.<sup>22</sup>

### Handicapped Children

"Handicapped children" includes those who are mentally retarded, physically handicapped, emotionally disturbed, learning disabled, speech impaired, hearing impaired, multiple handicapped or otherwise handicapped as defined by the Board of Education.<sup>23</sup>

### Hearing Impaired

Children who are hearing impaired are those children whose hearing loss (after all necessary medical treatment, surgery, and/or use of hearing aids) significantly restricts benefit from or participation in a normal classroom program and necessitates a modified instructional program.<sup>24</sup>

### Learning Disabled

Children with special learning disabilities exhibit a disorder in one or more of the basic psychological processes involved in understanding or in using spoken or written languages. These may be disorders of listening, thinking, reading, writing, spelling, or arithmetic. They include perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, or developmental aphasia. Learning problems which are due primarily to visual, hearing or motor handicaps, to mental retardation, emotional disturbance or to environmental disadvantage are not included.<sup>25</sup>

---

<sup>22</sup>Education Section, Division of Special Education, Administrative Requirements and Guidelines for Special Education Programs (Richmond, Virginia: State Department of Education, May 1972), p. 2.

<sup>23</sup>Education Section, Division of Special Education, p. 2.

<sup>24</sup>Education Section, Division of Special Education, p. 3.

<sup>25</sup>Education Section, Division of Special Education, p. 3.

### Mentally Retarded

Mentally retarded children are children whose mental capacity is such that they cannot be adequately educated in the regular classes in the public schools.

(a) Educable mentally retarded children are those who reveal a reduced rate of intellectual development and level of academic achievement below that of their peer age group as evidenced by significant deficits in all essential learning processes. . . .

(b) Trainable mentally retarded children are those whose educational needs cannot be met in a program designed for the educable mentally retarded because of an inability to acquire necessary skills as determined by a substantially reduced rate of intellectual development.<sup>26</sup>

### Multiple Handicapped

Children who are multiple handicapped are those whose combination of severely handicapping conditions requires extraordinary programs and/or services to meet their particular educational needs.<sup>27</sup>

### Physically Handicapped

Children who are physically handicapped are those whose physical condition(s) and/or special health problems result in the need for special provisions for educational purposes. This includes those children with organic, muscular, and neurological conditions affecting motor activities.<sup>28</sup>

### Special Education

"Special education" means classroom, home, hospital, institutional or other instruction to meet the needs of handicapped children; transportation, and corrective and supporting services required to assist handicapped children in taking advantage of, or responding to, educational programs and opportunities.<sup>29</sup>

---

<sup>26</sup> Education Section, Division of Special Education, p. 2.

<sup>27</sup> Education Section, Division of Special Education, p. 3.

<sup>28</sup> Education Section, Division of Special Education, p. 2.

<sup>29</sup> Education Section, Division of Special Education, p. 3.

### Speech Impaired

Children who are speech impaired have abnormality of speech which calls adverse attention to itself or interferes with communication which may be related to problems with articulation, rhythm, voice, and/or oral language.<sup>30</sup>

### Visually Impaired

(a) Children who are legally blind have 20/200 vision or less in the better eye with the best correction. . . .

(b) Children who are partially sighted 20/70 vision or less in the better eye after best correction, up to but not including the definition for legal blindness. In some instances an eye doctor may recommend for services a child who has better visual acuity than 20/70.<sup>31</sup>

### Limitations

This study included the following limitations:

1. The population from which the sample was drawn was from a school division in the State of Virginia located near Charlottesville, Virginia, and was limited to no more than the first 200 students in grades one through five and no more than the first 100 students in grades six through twelve referred to Special Services by their teachers, guidance counselor, or principal.
2. The cut-off period for the referrals was limited to March 31, 1978, or earlier if the required sample number for grades one through five and six through twelve had been attained.
3. The predictor was limited to the following tests:
  - A. SIT
  - B. SBIS

---

<sup>30</sup> Education Section, Division of Special Education, p. 3.

<sup>31</sup> Education Section, Division of Special Education, p. 3.

### C. WISC-R

4. The criterion was limited to the PIAT test.
5. The time schedule limits for the administration of the three selected mental maturity measures was from October 1977 through March 1978. These tests were given within a three-week period from the time of referral.
6. The administration of the PIAT was scheduled during May of 1978.
7. The level of significance in this study was .05.
8. The statistical test used in this study was limited to the Pearson Product-Moment Correlation.

### Significance of the Study

The significance of this study which included background, justification, and sources of data was as follows:

#### Background

Instruction was a primary responsibility of administration. This responsibility implied the following needs:

1. Increased accuracy in decision making;
2. Implemented evaluation strategies of greater depth and quality in relating to the student's needs as a process of prediction;
3. Sustained accountability through educational goals and outcomes.

The primary function of the administrator was to make decisions. Because the area of Pupil Personnel Administration involved delegated responsibilities in the areas of referral processes, identification,

classification, observation, diagnosis, consultation, evaluation, testing, educational recommendations, and implementation of instructional prescriptions, it was imperative that the responsible administrator increase the accuracy of his decisions, that his evaluation strategies were implemented with a depth and quality that related to the student's needs as a process of prediction, and that the effectiveness of instructional and administrative strategies provided sustained educational accountability.

### Justification

As an exercise in educational decision making, the first demand on the educational administrator was to make decisions with a high degree of accuracy.

Intelligence tests and achievement tests were constantly used by public school officials as aids in the decision-making process. "An investigator studies predictive validity when his primary interest is in bettering some outcome. The outcome is what we want to improve by our professional decisions."<sup>32</sup>

Sound decisions arise out of relevant knowledge of the individual . . . the more we know about a person that relates to our present decision, and the more accurately we know it, the more likely we are to arrive at a sound decision about him or a wise plan of action for him. . . .<sup>33</sup>

The second demand on the educational administrator was to implement evaluation strategies of greater depth and quality in relating

---

<sup>32</sup>Lee J. Cronbach, Essentials of Psychological Testing (3d ed.; New York: Harper and Row, 1970), p. 126.

<sup>33</sup>Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education (3d ed.; New York: John Wiley and Sons, 1969), p. 8.

to the student's needs as a process of prediction.

The evaluation of intelligence can be viewed as a process of prediction . . . the evaluation is used to predict the child's present level of functioning in problem solving situations, to predict future level of functioning, and to predict differences in functioning according to variations in internal and external circumstances.<sup>34</sup>

Research indicates that tests are among the best predictions available.<sup>35</sup>

P. E. Vernon summed it up when he said

. . . an intelligence test gives a better estimate of potentiality than other measures of achievement . . . its main usefulness lies in its ability to predict educability or trainability because of its greater generality and because it samples the reasoning capacities developed outside school which the child should be able to apply in school, e.g., to new subjects.<sup>36</sup>

The third demand on the educational administrator was to sustain accountability through educational goals by way of instructional and administrative strategies as well as in the assessment of educational outcomes achieved through its programs and activities.

Since 1970 the literature had been saturated with articles, books, and speeches on accountability in education. Much of this was due to new legislation, known as standards of quality, adopted by State Boards of Education mandating performance, professional, and system accountability. In addition, P.L. 94-142 was a national mandate for school systems to provide each handicapped child in need of special education and related

---

<sup>34</sup> Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), p. 253.

<sup>35</sup> Division of Special Services, Guidance Handbook for Virginia Schools, XLVII, No. 11 (Richmond, Virginia: State Department of Education, June 1965), p. 97.

<sup>36</sup> P. E. Vernon, Intelligence and Cultural Environment (London: Methuen, 1969), p. 27.



services a written individualized educational program. Statements concerning accountability such as "that the courts were laying new bases for judging teacher competence . . . results via student learning;"<sup>37</sup> the criterion of teacher effectiveness was formulated in "terms of the results teachers are able to produce in students."<sup>38</sup> Opposition to tests from educators had been viewed with "suspicion by the public . . . unless educators can develop more valid and dependable measures of pupil achievement than tests provide, the use of tests is not likely to diminish."<sup>39</sup> Obviously, the passing of P.L. 94-142 "requires teaching accountability . . . any arguments about the need for accountability are now moot. It appears to be the law of the land."<sup>40</sup>

The child's level as measured by an intelligence test provides one of the best clues available to the teacher as to the child's potentialities for learning . . . a guide as to what can reasonably be expected of each pupil.<sup>41</sup>

Public demands for educational accountability "require the proper use of . . . achievement tests to assess the results of the educational process."<sup>42</sup>

---

<sup>37</sup> C. H. Johnson, Jr., "Court, Craft, and Competence: A Reexamination of Teacher Evaluation Procedures," Phi Delta Kappan, IX (1976), 606.

<sup>38</sup> C. Knudsen, Evaluation and Improvement of Teaching (New York: Doubleday, Doran, 1932), p. 19.

<sup>39</sup> R. Ebel, "Declining Scores: A Conservative Explanation," Phi Delta Kappan, IV (1976), 309.

<sup>40</sup> Carlene Van Etten and Alen Van Etten, "The Measurement of Pupil Progress and Selecting Instructional Materials," Journal of Learning Disabilities, IX (October, 1976), 4.

<sup>41</sup> Thorndike and Hagen, pp. 338-39.

<sup>42</sup> Anne Anastasi, Psychological Testing (4th ed.; New York: Macmillan, 1976), p. 403.

### Sources of Data

The primary sources of data were the test results gained from each of the referrals.

### Assumptions

The assumptions pertinent to this study were as follows:

1. Test scores were used as aids in decision making and as sources of data on which to base further investigation.
2. Tests measured important behaviors.
3. Intelligence and achievement tests combined gave better predictions to later school achievement.
4. Sound decisions arose out of relevant knowledge of the individual.
5. Subjects being tested had been exposed to comparable, but not necessarily identical, acculturation.
6. The specific function of administration was to develop and regulate the decision-making process in the most effective manner possible.

### Hypotheses

The primary hypothesis and sub-hypotheses related to this study were as follows:

#### A. Primary Hypothesis

There was no significant difference in the predictive ability between the SIT, SBIS, and the WISC-R, when compared to academic achievement in mathematics and reading as measured on the PIAT.

## B. Sub-Hypotheses

There were no significant differences between mental ages, as derived from SIT, SBIS, and the WISC-R, and the MA as derived from grade placement on mathematics and reading of the PIAT.

1. There was no significant difference between the SIT MA and the Reading MA.
2. . . . between the SIT MA and the Mathematics MA.
3. . . . between the SBIS MA and the Reading MA.
4. . . . between the SBIS MA and the Mathematics MA.
5. . . . between the WISC-R(V) MA and the Reading MA.
6. . . . between the WISC-R(V) MA and the Mathematics MA.
7. . . . between the WISC-R(P) MA and the Reading MA.
8. . . . between the WISC-R(P) MA and the Mathematics MA.
9. . . . between the WISC-R(FS) MA and the Reading MA.
10. . . . between the WISC-R(FS) MA and the Mathematics MA.

## Procedures

The procedures in this study were as follows:

## Related Studies

Correlations between scores on intelligence tests and academic achievement generally fell in the .40-.70 range. Correlations were generally higher if the criteria were based on a standardized measure of achievement and that criterion was more academically verbal in content such as English, reading, or mathematics.<sup>43</sup>

---

<sup>43</sup> Frederick Brown, Principles of Educational and Psychological Testing (Hinsdale, Illinois: The Dryden Press, 1970), p. 336.

When predicting academic performance from scholastic aptitude test scores, the validity coefficients generally cluster around .50-.60 . . . this level of predictive accuracy has been essentially constant over the past several decades, even when new tests or analytic methods have been introduced.<sup>44</sup>

Correlations were

. . . higher in elementary schools than in high schools and in high schools than in colleges . . . studies in the past have indicated a drop in correlation from .70 in elementary school to .60 in high school and .50 in college . . . the drop in correlation is probably to be explained by the decreased range of intellectual ability in the college groups.<sup>45</sup>

It appeared that the most common use of tests, other than classroom examinations, was for the prediction of some future behavior.<sup>46</sup> However, "studies of the predictive validity of individual tests on representative samples are scarce because the tests are ordinarily applied only to cases referred for special study."<sup>47</sup> The sample for the study was drawn from a special population; all students were presumed to have a handicapped condition, thus, the sample was very appropriate for a predictive validity study.

Reports from the literature are extensively treated in Chapter 2. However, some of the more important studies that relate to this study in format and content were as follows:

1. The comprehension test of the Gates-MacGinitie Reading Tests,

---

<sup>44</sup> Brown, p. 431.

<sup>45</sup> Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education (3d ed.; New York: John Wiley and Sons, 1969), p. 324.

<sup>46</sup> Brown, p. 102.

<sup>47</sup> Lee J. Cronbach, Essentials of Psychological Testing (3d ed.; New York: Harper and Row, 1970), p. 235.

Survey F for grades ten through twelve, correlated highly (.79) with the Lorge-Thorndike Verbal IQ scores. The test had an alternate-form reliability of about .88.<sup>48</sup>

2. The median correlation between the Otis IQ and the Metropolitan Achievement High School Battery Tests for grades nine through thirteen was .68. This correlation included the ten subtests of the achievement battery excluding the Language Study Skills section.<sup>49</sup>

3. The Peabody Individual Achievement Test for grades kindergarten through twelve, showed median correlations of .68 with Peabody Picture Vocabulary Test IQ's. This test was very helpful in evaluating pupils who may be in need of special study. Total test scores were a reflection of overall school achievement. The median test-retest reliability for the Peabody Individual Achievement Test was .89 for the total test.<sup>50</sup>

4. A study which used the Lorge-Thorndike Intelligence Test (verbal) as a predictor and the Iowa Test of Basic Skills, grade four, as a criterion, showed the validity coefficients to be .78. Another study using the American College Testing Program Test Index as a predictor and college grades in English and math as a criterion, revealed a validity coefficient of .54 for English grades and .44 for mathematics grades.

Intelligence tests correlated higher with standardized measures of achievement than with school marks. Correlations between an

---

<sup>48</sup> Oscar Buros, ed., The Seventh Mental Measurements Yearbook, Volumes I and II (Highland Park, New Jersey: Gryphon Press, 1972), p. 1084.

<sup>49</sup> Buros, Volume I, p. 32.

<sup>50</sup> Buros, Volume I, pp. 34-35.

intelligence test and total score on an achievement battery in the .70's or even .80's were not unusual. In a study using a sample of 2500 students in each grade, the following correlations were highly significant for each sub-test area using the Lorge-Thorndike Intelligence Test as a predictor and the Iowa Tests of Basic Skills as a criterion. The correlations were as follows:<sup>51</sup>

	<u>L-T Verbal</u>	<u>L-T Nonverbal</u>
Iowa Vocabulary	.71-.82	.56-.65
Iowa Reading	.68-.82	.53-.69
Iowa Language	.73-.79	.61-.67
Iowa Study Skills	.72-.81	.62-.78
Iowa Arithmetic	.66-.75	.61-.71
Iowa Composite	.79-.88	.65-.77

5. Some of the more important tests used for screening devices for follow-up evaluations and for assessing handicapped as well as normal children revealed significant validity correlations. In addition, these tests were designed to assess handicapped children. A screening procedure was needed when time was limited, when qualified psychologists were not available, and when a handicapped child was not able to perform due to verbal and motor handicaps.<sup>52</sup>

The validity correlations were as follows:<sup>53</sup>

---

<sup>51</sup>Thorndike and Hagen, pp. 170, 324.

<sup>52</sup>Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), pp. 235-36.

<sup>53</sup>Sattler, pp. 236-46, 429.

## a. Peabody Picture Vocabulary Test (PPVT) (1959, 1965)

<u>Binet (1960)</u>	<u>Wechsler (1949)</u>
.66 (median)	.66 (median) VIQ
	.63 (median) FSIQ
	.54 (median) PIQ

## b. Quick Test (IQ) (1962)

Iowa Test of  
Basic Skills

<u>PPVT (1965)</u>	<u>(1955-56)</u>	<u>Binet (1960)</u>	<u>Wechsler (1949)</u>
.76 (median)	.48 (median)	.61 (median)	.31-.88 VIQ
			.22-.70 PIQ
			.35-.84 FSIQ

## c. Pictorial Test of Intelligence

<u>Binet (1960)</u>	<u>Wechsler (1949)</u>
.72	.65 (FSIQ)
<u>Columbia Mental Maturity Scale (1972)</u>	<u>Wide Range Achievement Test (1965)</u>
.53	.56 (Reading)
	.79 (Arithmetic)

## d. Columbia Mental Maturity Scale

<u>Binet (1960)</u>	<u>Wechsler Scales (all) (1949)</u>
.65 (using mental ages)	.64 (median)
.56 (using IQ)	

Treatment of the Data

The following is an outline of procedures as it relates to the treatment of data.

1. There was a comprehensive search of the literature (Counseling and Personnel Services-Direct Access to Reference Information: A Xerox

Service-Mental Measurement Yearbooks-Journals-Periodicals-Manuals).

2. Criteria were established for the purpose of selecting the three mental maturity tests and the one achievement test to be used in this study.

3. The sample was limited to no more than the first 200 students in grades one through five and no more than the first 100 students in grades six through twelve referred to special services by their teachers, guidance counselor or principal. All of the students referred were presumed to have some type of handicapped condition. Approval was given by the Albemarle County School System relative to using human subjects.

4. The entire sample was tested using the SIT, SBIS, and the WISC-R. The PIAT was administered to the entire sample during May of 1978.

5. The primary and sub-hypotheses were tested in the null format.

6. The level of significance in this study was .05. The statistical test used was the Pearson Product-Moment correlation. The resulting  $\gamma$  statistic was the validity coefficient  $\gamma_{xy}$  which was the correlation between predictor and criterion. The higher the validity coefficient, the greater was the correlation between the two variables.

7. The summary included findings, conclusions, and recommendations.

### Summary

This study is organized in the following manner:

Chapter 1 includes the introduction, defines the problem, defines terms, establishes the limitations of the study, explains the significance of the study, develops assumptions and hypotheses, explains the procedures of the research, and summarizes the organization of the dissertation.



Literature related to the study is presented in Chapter 2.

Chapter 3 includes the methodology used in the research design, the procedures executed to collect data, and the statistical treatment utilized to analyze the data.

Chapter 4 is an analysis of the research.

Chapter 5 is a summary of the research findings, conclusions, and recommendations.

## Chapter 2

### REVIEW OF RELATED LITERATURE

#### Introduction

The importance of decision making in administration had been recognized since the writing of Barnard in 1938, Simon in 1947, and Bross in 1953.<sup>1</sup> "Decision-making was becoming generally recognized as the heart of organization and the process of administration."<sup>2</sup> James McCammy suggested that the making of decisions was the center of the administrative process.<sup>3</sup>

Intelligence tests and achievement tests had been constantly used as aids in the decision-making process. "An investigator studies predictive validity when his primary interest is in bettering some outcome. The outcome is what we want to improve by our professional decisions."<sup>4</sup>

Sound decisions arise out of relevant knowledge of the individual . . . the more we know about a person that relates to our present decision, and the more accurately

---

<sup>1</sup>Stephen J. Knezevich, Administration of Public Education (3d ed.; New York: Harper and Row, 1975), p. 59.

<sup>2</sup>Daniel E. Griffiths, Administrative Theory (Englewood Cliffs, New Jersey: Prentice-Hall, 1959), p. 75.

<sup>3</sup>James L. McCammy, "Analysis of the Process of Decision Making," Public Administration Review, VII (1947), 41.

<sup>4</sup>Lee J. Cronbach, Essentials of Psychological Testing (3d ed.; New York: Harper and Row, 1970), p. 126.

we know it, the more likely we are to arrive at a sound decision about him or a wise plan of action for him. . . .<sup>5</sup>

The outstanding success of scientific measurement of individual differences has been that of the general mental test.<sup>6</sup> Despite the frequent criticisms leveled against the IQ, a child's IQ, obtained in a standard situation, has more demonstrated behavioral correlates than any other psychological measure . . . individual intelligence tests adequately predict scholastic achievement, yield a more useful picture of cognitive development than group tests, and aid in clinical situations.<sup>7</sup>

"The adequacy of the entire decision-making process hinges more on the adequacy of the criteria than on any other single aspect of the situation."<sup>8</sup>

Among the criteria most frequently employed in validating intelligence tests is some index of academic achievement . . . the various indices of academic achievement have provided criterion data at all educational levels . . . employed principally in the validation of general intelligence tests.<sup>9</sup>

Instruction was a primary responsibility of administration, and it was imperative that valid and reliable standardized instruments be used in the prediction of one's academic level. When valid and reliable prediction instruments were used, the accuracy of instructional and administrative decision-making was enhanced. It is imperative that the

---

<sup>5</sup>Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education (3d ed.; New York: John Wiley and Sons, 1969), p. 8.

<sup>6</sup>Cronbach, p. 197.

<sup>7</sup>Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), p. 23.

<sup>8</sup>Frederick Brown, Principles of Educational and Psychological Testing (Hinsdale, Illinois: The Dryden Press, 1970), p. 105.

<sup>9</sup>Anne Anastasi, Psychological Testing (4th ed.; New York: Macmillan, 1976), p. 142.

educational prescription be instructionally tailored and commensurate with each child's abilities and disabilities.

Individual differences among students is the most persistent and baffling problem in education today. The area of most significance involves the student's ability to achieve . . . there is no denying that intelligence is one of the crucial elements in ability to master the curriculum of studies.<sup>10</sup>

### Administrative Decision-Making

The recurrent theme in administrative decision-making was that in any organizational administrative operation the making of decisions was at the center of all administrative processes.

The need for administration was due to the increased complexities of the educational institutions where man had "to organize and manage his resources via specialized institutions to attain educational, political, economic, and social goals."<sup>11</sup>

One of the most important contributions administration had in the operation of an institutional enterprise was the implementation of decisions.

The anatomy of the organization is to be found in the distribution and allocation of decision-making functions. The physiology of the organization is to be found in the processes whereby the organization influences the decisions of each of its members supplying these decisions with their premises.<sup>12</sup>

---

<sup>10</sup>Max Wingo, Philosophies of Education: An Introduction (Lexington, Massachusetts: D. C. Heath, 1974), p. 73.

<sup>11</sup>Knezevich, p. 3.

<sup>12</sup>Herbert A. Simon, Administrative Behavior (2d ed.; New York: Macmillan, 1957), p. 220.

"Decisioning in an organization is not a personal matter, and the effectiveness of decisions is not a product of the quality of decisions of any one person."<sup>13</sup> Livingston suggested that the concept of decision making was not only the process by which the decision was arrived at, but it also incorporated the process by which the decision was implemented.<sup>14</sup>

#### The decision process

. . . is an organizational matter, and the criterion by which an organization may be evaluated is the quality of the decisions which the organization makes plus the efficiency with which the organization puts the decisions into effect.<sup>15</sup>

#### The specific function of administration

. . . is to develop and regulate the decision-making process in the most effective manner possible . . . effective manner in one which results in the accomplishment of a stated objective.<sup>16</sup>

Therefore, any decision was a judgment which affected some type of action.<sup>17</sup>

There existed the belief that decisions of all types were made by a succession of steps. Since the process of decision making was a cycle of events where a consistent quality or direction could be discerned, the following steps were necessary for implementation:<sup>18</sup>

The first formal step in the decision-making process was to recognize there existed a problem and the decision-maker had gone through a process of defining and limiting the problem.<sup>19</sup>

---

<sup>13</sup>Griffiths, p. 113.

<sup>14</sup>Robert Teviot Livingston, "The Theory of Organization and Management," Transactions of the ASME, May, 1953, p. 659.

<sup>15</sup>Griffiths, p. 113.

<sup>16</sup>Griffiths, p. 73.

<sup>17</sup>Griffiths, p. 76.

<sup>18</sup>Griffiths, pp. 92-94.

<sup>19</sup>Griffiths, p. 95.

The second formal step in the decision-making process was to analyze and evaluate the problem. The decision-maker needed to decide whether or not he would attempt to solve it.<sup>20</sup>

The third formal step was to set up established criteria whereby solutions would be evaluated as being acceptable and adequate to the need.<sup>21</sup> This was a crucial stage because individualized values and goals of the organization were built into the process.

The fourth formal step was to collect relevant data based on the decision to be made.<sup>22</sup>

The fifth formal step was the formulation and selection of preferred solution(s).<sup>23</sup> At this phase of the process there were occurring weighted consequences of each solution.

The sixth formal step was the implementation of the preferred solution(s). Included in this last step were three phases of the implemented solution related to programming the solution, controlling the activity, and evaluating the results.<sup>24</sup>

Administrative decisions were those which establish criteria for others in the organization to make their decisions. In the words of the engineer, "an administrator is a type of servo mechanism, establishing the limits within which function is controlled."<sup>25</sup> Therefore, the quality of any decision was largely determined by the administrator's decisions.

---

<sup>20</sup>Daniel E. Griffiths, Administrative Theory (Englewood Cliffs, New Jersey: Prentice-Hall, 1959), p. 97.

<sup>21</sup>Griffiths, p. 102.

<sup>22</sup>Griffiths, p. 103.

<sup>23</sup>Griffiths, p. 104.

<sup>24</sup>Griffiths, p. 107.

<sup>25</sup>Griffiths, p. 93.

### Nature of Intelligence and Psychological Testing

Intelligence testing was responsible for bringing psychology into being as a separate discipline. "Intelligence testing had its roots in the fields of general psychology and measurement."<sup>26</sup>

The first systematic experimentation on individual differences in behavior had come from the discovery that astronomers differed in reaction time. In 1796, an assistant named Kinnebrook at Greenwich Observatory "was engaged in recording, with great precision, the instant when certain stars crossed the field of the telescope."<sup>27</sup> When Kinnebrook consistently reported observations different from those of his supervisor, he was discharged for incompetence.

In 1916 Bessell,

Astronomer at Konigsberg, read of the Kinnebrook incident . . . and decided to look further into such observational errors . . . Bessell set out to discover whether such personal differences could be found among more experienced astronomical observers.<sup>28</sup>

Astronomers had also become interested in other conditions which affected the magnitude of personal error such as visual-versus-auditory modality and the rate of movement of the stimulus.<sup>29</sup>

In 1879, Wilhelm Wundt established the first laboratory of experimental psychology at Leipzig. These experimental psychologists were trained chiefly in physics and physiology and it was characteristic

---

<sup>26</sup>Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), p. 7.

<sup>27</sup>Lee J. Cronbach, Essentials of Psychological Testing (3d ed.; New York: Harper and Row, 1970), p. 126.

<sup>28</sup>Anne Anastasi, Individual Differences (New York: John Wiley and Sons, 1965), p. 2.

<sup>29</sup>Anastasi, Individual Differences, p. 2.

of the period for experimental psychologists to either ignore individual differences or to treat them simply as chance errors. Therefore, the attitude toward individual differences was regarded as the "margin of error to be expected in the application of general laws of psychology."<sup>30</sup> However, one important discovery did come from the rise of experimental psychology: the controlling of extraneous variables in measuring individual behavior, as well as the standardization of testing materials and procedures.<sup>31</sup>

Sir Francis Galton, an English biologist, whose concern for individual differences helped set up a psychometric laboratory at the International Health Exhibition in 1884, later known as University College, London. Galton's efforts had been very active in the field of mental measurement, in the study of the inheritance of intellectual ability and the development of statistical methods. Galton saw the measurement of sensory capacities as a promising method of gauging the intellectual level. Galton wrote,

The only information that reaches us concerning outward events appears to pass through the avenue of our senses; and the more perceptive the senses are of differences, the larger is the field upon which our judgment and intelligence can act.<sup>32</sup>

Galton was also responsible for the application of rating-scales and questionnaire methods. In addition to his work with sensorimotor tests, Galton had extended enormously the application of statistical procedures

---

<sup>30</sup> Anastasi, Individual Differences, p. 3.

<sup>31</sup> Anne Anastasi, Psychological Testing (4th ed.; New York: Macmillan, 1976), p. 7.

<sup>32</sup> Francis Galton, Inquiries into Human Faculty and Its Development (London: Macmillan, 1883), p. 27.



to the analysis of test data. One of his students, the eminent Karl Pearson, was credited with following Galton's work in developing the details of correlation theory and correlation coefficients.

Another outstanding contributor to the development of psychological testing was the American psychologist, James Cattell. Cattell was known in America for his psychological laboratories and the development of the testing movement. Cattell

. . . introduced the term "mental test" in one of his published articles in 1890 called Mental Tests and Measurements . . . the essence of this article describes a series of tests that were being administered annually to college students in an effort to measure their intellectual level . . . Cattell shared Galton's view that an estimate of intellectual functioning could be obtained through tests of sensory discrimination and reaction time . . . it was further supported by his conviction that simple functions could be measured with precision, while the more complex, that of higher mental processes, could not.<sup>33</sup>

At the same time other individuals in the United States such as Jastrow of the University of Wisconsin, Boas at Clark University, and Gilbert of New Haven were demonstrating how children responded to various types of tests. In addition to these men, in France Binet, Henri, and Simon were developing methods for the study of a variety of mental functions. The key to the measurement of intelligence for Binet was focusing on higher mental processes instead of simple sensory functions.<sup>34</sup> In many of these functions such as memory, attention, judgment, reasoning, comprehension, and imagination one recognized the forerunners of the famous Binet intelligence tests.

---

<sup>33</sup>Anastasi, Individual Differences, p. 6.

<sup>34</sup>Sattler, p. 8.

In 1904, the French Minister of Public Instruction appointed a commission to study the problem of retardation among public school children. Out of his work for this commission, Binet, in collaboration with Simon, "prepared the first intelligence scale designed to yield a global index of intellectual level . . . a child's score on the scale could then be expressed as a mental age."<sup>35</sup> Translations and adaptations had appeared in many countries, and in America several revisions had been prepared, including the well-known Stanford-Binet by Terman and his associates at Stanford University.

Another important milestone in the mental testing movement was the development of group tests. The Binet scales were individual tests requiring a highly trained examiner to administer and interpret them. They were not suited for large-scale testing and were used as clinical instruments for intensive study of individual cases. Group intelligence tests had been a major factor in the popularization of psychological testing. Given to large groups at one time, they were easy to administer and score. The stimulus for the development of group tests was provided by the entrance of the United States into World War I in 1917. The American Psychological Association met to consider ways in which psychology could help in the conduct of the war. The result of their study was the utilization of Army Alpha and Army Beta tests to classify recruits with respect to their general intellectual level. Administrative decisions were based on intellectual scores for such decisions as rejection or discharge from military service, assignments to different types of service, and entrance to officer

---

<sup>35</sup> Anastasi, Individual Differences, p. 7.

training camps. However, around 1920 studies of the intellectual performance of individuals began to focus on the effects of early experience of such groups as the culturally deprived, orphanages, gypsy camps, isolated mountain communities and city slums. After the termination of World War I, the "Army tests were released for civilian use . . . they served as models for most group intelligence tests."<sup>36</sup>

In 1927 Charles Spearman discovered the factor analytic approach to intelligence. This approach to intelligence was known as the two-factor theory where a general factor (g) plus a specific factor (s) per test accounted for an individual's performance on an intelligence test. Any intellectual activity "involves a general factor which it shares with all other intellectual activities, and a specific factor which it shares with none."<sup>37</sup>

In 1927 Thorndike purposed that intelligence was based on a multitude of separate elements, each representing a distinct ability. Certain mental activities had elements in common and combined to form clusters. Thorndike identified his clusters as social intelligence (handling people), concrete intelligence (handling things), and abstract intelligence (handling verbal and mathematical symbols).

In 1938 Thurstone claimed that intelligence was a small set of group factors which are "primary mental abilities . . . these abilities are verbal meaning, number facility, inductive reasoning, perceptual speed, spatial relations, memory, and verbal fluency."<sup>38</sup>

---

<sup>36</sup> Anastasi, Psychological Testing, p. 13.

<sup>37</sup> Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), p. 10.

<sup>38</sup> Sattler, p. 10.

As early as 1919 Burt showed evidence for the increasing differentiation of intellectual ability with age and he agreed with Garrett that the differentiation of abilities was due chiefly to maturation. In 1955 Burt observed that mental capacity was cognitive, general, and innate.

In 1950 Vernon developed a hierarchical theory of intelligence. The highest level was a general intellectual factor (g), followed by two major group factors--verbal-educational, and practical-mechanical-spatial. This "theory synthesizes the work of Spearman and Thurstone, but gives central importance to the general factor (g)."<sup>39</sup> Prior to 1950, in fact 1946, Garrett suggested that intelligence was developmental. Garrett sums it up by saying

greater differentiation appears at the upper age levels. . . .<sup>40</sup>

. . . We can predict a steady drop in correlation among tests involving verbal, numerical, and spatial concepts from about age 8 to age 18. . . .<sup>41</sup>

. . . That the (g) factor is strongest at the elementary school level and is in large part, verbal or linguistic in nature . . . if the school child can read well he can very probably do the rest of his school work well. Solving arithmetic problems is contingent upon ability to read and understand directions; hence a fifth grade child high in verbal facility may do as well in arithmetic as a child of much greater native aptitude for numbers. . . .<sup>42</sup>

. . . The total score on an individual intelligence test is a better measure of general ability than part scores on the same test can possibly be measures for more specific functions.<sup>43</sup>

---

<sup>39</sup>Sattler, p. 11.

<sup>40</sup>Anne Anastasi, Individual Differences (New York: John Wiley and Sons, 1965), p. 77.

<sup>41</sup>Anastasi, Individual Differences, p. 79.

<sup>42</sup>Anastasi, Individual Differences, p. 80.

<sup>43</sup>Anastasi, Individual Differences, p. 82.

In 1967 Guilford, a prominent multifactor theorist, showed the organization of intellectual factors by developing a structure of intellect model. The model was three dimensional with one dimension representing operation categories, the second dimension representing content categories, and a third dimension representing product categories. The model had five operations (cognition, memory, divergent thinking, convergent thinking, and evaluation), four types of content (figure, symbolic, semantic, and behavioral), and six products (units, classes, relations, systems, transformations, and implications). There were approximately 120 factors.

Wesman purposed that intelligence was an attribute not an entity; and intelligence was the summation of the learning experiences of the individual.

Jensen strongly suggested that intelligence tests were valuable tools which provide reliable and valid measures of abilities needed in contemporary society. Based on his experimental work of 1970, Jensen concurred that intelligence consisted of associative ability (memory and serial learning tasks), and cognitive ability (abstract reasoning tasks).

It was necessary to understand the meanings associated with intelligence, as well as the definitions of intelligence, in order to give greater insight into the theories of intelligence. Vernon suggests three meanings: (1) "intelligence is the innate capacity of the individual . . . (2) intelligence is what the individual does or his observed behavior . . . (3) intelligence is the result obtained on an intelligence test."<sup>44</sup>

---

<sup>44</sup>Sattler, p. 8.

Themes in the history of intelligence revealed a general progression from no accepted definition of intelligence or method of testing intelligence to a conception of intelligence based on logical or empirical approaches. At times approaches had been trial-and-error as well as intuitional, but as time passed these approaches had been superseded by logical, systematic, and empirical approaches.

One of the things that define intelligence is its correlations with some kinds of criteria . . . another important dimension is the factor analytic approach which is a statistical way of sorting out components which are related to overall task performance.<sup>45</sup>

Defining intelligence continued to be a problem in the sense of arriving at a common definition. Terman defined intelligence as abstract thinking; Binet defined intelligence as a collection of faculties; Wechsler was known to view intelligence as having the qualities of purposefulness, rationality, and ability to deal effectively with the environment; Piaget suggested that intelligence was a biological adaptive process of assimilation and accommodation. The writer took the view of Wechsler that intelligence was "the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment."<sup>46</sup> Theories of intelligence were beginning to emerge showing a coalescent view, which stressed the importance of innate and developmental trends.

#### Measurement of Intelligence

An evaluation of the constancy of the IQ requires consideration

---

<sup>45</sup>Sattler, p. 6.

<sup>46</sup>David Wechsler, The Measurement and Appraisal of Adult Intelligence (4th ed.; Baltimore: Williams and Wilkins, 1958), p. 7.

of prediction and regularity of intellectual development. Prediction was good for short periods of time, but the IQ fails to give a stable index of development due to large shifts over an extended time interval. Within the first four years of life as much development takes place as in the next thirteen years.

The family environment had affected the rate of mental growth as had personality patterns in children. Socioeconomic factors were also related to intelligence. L. E. Tyler says that "the relationship of measured intelligence to socio-economic level is one of the best documented findings in mental-test history."<sup>47</sup> The relationship between these factors and the theories, definitions, and meanings of intelligence suggested that some of the limitations of intelligence testing and intelligence tests were associated with predicting occupational success and nonacademic skills, providing measures of minute capacity, cognitive functions, and processes underlying test responses, and possible penalizing for nonconventional responses as well as possible unreliability for excessive long range predictions. However, despite the accusations concerning the limitations of intelligence testing and intelligence tests, the IQ obtained on a standard intelligence test had better demonstrated behavioral correlates than any other psychological measure. Tests provided the measurement of change, related information about the individual, and assisted in understanding the variables relative to the nature of intelligence and environment.<sup>48</sup>

---

<sup>47</sup> L. E. Tyler, The Psychology of Human Differences (3d ed.; New York: Appleton-Century-Crofts, 1965), p. 336.

<sup>48</sup> Anne Anastasi, "Psychology, Psychologists, and Psychological Testing," American Psychologists, XXII (1967), 297-306.

Other issues which involved the measurement of intelligence were selection of tests to be used and the value of intelligence tests.

The SBIS, WISC-R, and SIT as screening instruments held the most promise for obtaining the fullest amount of information about the child, while also providing a measure of intelligence. The SBIS and WISC/WISC-R were highly preferred in hospitals and clinics and in institutions for the mentally retarded. The SBIS was preferred for giftedness and for mental retardation in kindergarten through second grades. Starting with third grade the WISC/WISC-R was preferred for mental retardation. After sixth grade, the WISC/WISC-R was preferred for evaluating problems concerned with differential diagnosis (e.g., learning problems, emotional problems, and neurological problems). Evidence also indicated that both the SBIS and the WISC/WISC-R had good predictive validity and were valid and reliable instruments used in predicting educational achievement as well as aids to the decision-making process.<sup>49</sup>

The value of intelligence tests was unlimited. Valuable information had been obtained in a relatively brief period of time, halo effects were reduced, a record was provided to make predictions and compare past scores, and excellent diagnostic and assessment information was provided.

### Intelligence Tests

#### Stanford-Binet Intelligence Scale

Alfred Binet was considered to be the father of intelligence testing. His areas of interests were developmental, clinical, and

---

<sup>49</sup> Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), pp. 414-15.



experimental psychology. Binet and a physician, Theodore Simon, did an extensive amount of work in the area of intellectual measurement, leading to the development of the Binet-Simon scales. Incentive for the construction of the first scale was provided by the Minister of Public Instruction in Paris in 1904, when a committee was appointed to find a method to separate the subnormal (mentally retarded) from the normal child in the schools. The 1905 scale by Binet-Simon was the result; it was based on a practical need of the public school system in Paris. The Binet-Simon scales from inception were recognized as being extremely valuable for the diagnosis of mental retardation.

The object of the 1905 scale had been to devise a measure of the intellectual capacities of school children. This scale

. . . measured general mental development and judgment rather than an assortment of specific functions . . . 30 tests were devised that included items of simple commands; coordination of movement of head and eye; tactile and visual activities; verbal knowledge of objects; ability to define words; knowledge of pictures; designation of objects; and completion of sentences. . . . Binet and Simon considered judgment, comprehension, and reasoning as the essential parts of intelligence. The process of thinking was conceived of as the ability to adopt and maintain a given set, the ability to make adaptations, and the ability to criticize oneself.<sup>50</sup>

Burt's tribute to the Binet-Simon scales showed the range of Binet's and Simon's contribution to the testing movement when he stated,

As a provisional but practicable plan for testing mental deficiency, as a rough but intelligible method of interpreting the results, as a pioneer investigation of the general course of mental development, as a demonstration of the richness of the higher, more complex, and more ordinary mental processes, as a protest against the mere examination of acuity of sensation, of speed of reaction, or of anatomical peculiarities, as a means of interesting the teacher, the doctor, and the social worker in the

---

<sup>50</sup> Sattler, pp. 91, 105.

measurement of psychological capacities by psychological devices, as a prolific source of inspiration and suggestion, and, finally, as a stimulus to scientific discussion and inquiry, in these and many other ways the Binet scheme remains a marvel and a masterpiece.<sup>51</sup>

The Binet-Simon scales had been well received in the United States. Goddard was the leading proponent of these scales and did much of the work in translating, adapting, and standardizing them. Goddard introduced the 1905 scale to the United States in 1908, and two years later, the 1908 scale in 1910. The 1908 scale was standardized on two thousand American children. For many years this was the version most commonly used. The year level format of tests was introduced in the 1908 scale, which had fifty-nine tests. The 1911 scale showed a further refinement of the scales and increased the range to include an adult year-level while decreasing the number of tests to fifty-four. The 1908 and 1911 scales were the first and second revisions of the Binet-Simon scales of 1905.

Terman became interested in the practical and theoretical value of the Binet-Simon scales and in 1916

. . . standardized the Binet-Simon scales, added tests, revised others, changed methods of scoring and administration, introduced alternative tests . . . increased the number of tests from 54 to 90, used Stern's mental quotient concept of dividing mental age by chronological age . . . the 1916 Stanford-Binet ranged from year-level III to Superior Adult I and was the first revision of the originally devised Binet-Simon Scales.<sup>52</sup>

The second revision of the Stanford-Binet scales appeared in 1939. Terman and Merrill were responsible for this revision which

. . . extends the range from year-level II to Superior Adult III, increased the number of tests to 129 and has two forms L and

---

<sup>51</sup>Sattler, p. 93.

<sup>52</sup>Sattler, pp. 94, 106.

M . . . standardized on 3184 native-born whites, and showed reliability coefficients for the 1916 and 1937 revisions ranged from .98 for subjects with IQ's below 70 to .90 for subjects with IQ's above 129. The median correlation was .62 and validity coefficients indicated that the relationship between IQ and school success ranges from .40 to .50.<sup>53</sup>

The third revision of the Stanford-Binet scales appeared in 1960 known as the combined L-M revision taken from the two 1939 forms.

Instructions for test administration and scoring were improved, the formats for the test materials were redesigned, the IQ tables were extended from age 16 to 18 but the scale was not restandardized . . . instead a sample of 4,498 subjects who had taken the scale between 1950 and 1954 was used to check on changes in item difficulty . . . one of the most important developments in Form L-M is the replacement of the 1937 scale's IQ tables by Pinneau's deviation IQ for ages 2 through 18 normalized with a mean of 100 and a standard deviation of 16, and increased the number of tests to 142.<sup>54</sup>

In a study by Bradway, Thompson, and Cravens, subjects who had been tested as part of the 1937 standardization were re-tested after a ten-year interval and again twenty-five years after the original testing making the age range "26.5 to 32.3 (mean age 29.5) is expressed by a Pearsonian  $r$  of .59 . . . this compares favorably with the Pearsonian  $r$  of .65 found for the same group in the first follow-up after only ten years."<sup>55</sup>

The distribution of the 1937 standardization sample had provided the frame of reference for classifying IQ's. This was shown in Table 1, page 39.<sup>56</sup>

---

<sup>53</sup> Sattler, pp. 101, 102, 106.

<sup>54</sup> Sattler, pp. 104, 106.

<sup>55</sup> Lewis M. Terman and Maud A. Merrill, Manual for the Stanford-Binet Intelligence Scale (3d ed.; Boston: Houghton Mifflin, 1973), p. 17.

<sup>56</sup> Terman and Merrill, p. 18.

Table 1  
Distribution of the 1937 Standardization Group

IQ	Classification
140 - up	Very superior
120 - 139	Superior
110 - 119	High average
90 - 109	Normal or average
80 - 89	Low average
70 - 79	Borderline defective
69 - Down	Mentally defective

### Validity and Reliability

Evidence of the 1960 scale stemmed from the

Choice of items according to mental age on the 1937 scale, assures that the new scale is measuring the same thing that was measured by the original scale, increase in mental age from one age to the next checked with increase in percent passing from one chronological age to the next in both forms of the 1937 scale, and the choice of items was determined by their correlation with total score on each form.

The mean correlation for the 1960 scale is .66, and compares with a mean of .61 for all tests in both forms in the 1937 version. Comparing mean correlations in 1937 and 1960 of only those subtests used to make up Form L-M, the same relative variation appears. The mean 1937 correlation for comparable subtests was .62. Additional evidence for the high reliability of the Stanford-Binet is that both Form L and Form M have high biserial correlations.<sup>57</sup>

Verbal tests had a higher validity than non-verbal tests.

Verbal average for the L-M scale was .65. The 1937 average for the L-M verbal test was .63. Non-verbal test of the L-M scale correlated .58 with the total scale whereas in 1937 there was a biserial correlation of .51.

---

<sup>57</sup> Terman and Merrill, pp. 32, 33.

The most valid tests of intelligence which best predicted one's level of problem solving ability are Vocabulary, Abstract Words, Sentence Building, Similarities and Differences, Analogies, Sentence Completion, Verbal Absurdities, and Reasoning . . . in the 1960 L-M scale the eight best tests had an average correlation of .73 with the total scale whereas in 1937 they correlated .68 with the total scale.<sup>58</sup>

### Classification System

Classification systems had been used because examiners were interested in the detailed analysis of an examinee's performance. A classification system was a convenient way of describing the child's strengths and weaknesses by analyzing the subtests and classifying them under factors, abilities, or learning components. These categories had much validity for interpretive purposes.

Two important classification systems that had been widely used are Valett's and Sattler's. Valett's was as follows:

A. General Comprehension - the ability to conceptualize and integrate components into a meaningful total relationship.

B. Visual-Motor Ability - the ability to manipulate material in problem-solving situations that usually require integration of visual and motor skills.

C. Arithmetic Reasoning - the ability to make appropriate numerical associations and to deal with mental abstractions in problem-solving situations.

D. Memory and Concentration - the ability to attend and retain. Requires motivation and attention and usually measures degree of retention of test items.

---

<sup>58</sup>Terman and Merrill, p. 34.

E. Vocabulary and Verbal Fluency - the ability to correctly use words in association with concrete or abstract material; the understanding of words and verbal concepts; the quality and quantity of verbal expression.

F. Judgment and Reasoning - the ability to comprehend and respond appropriately in specific situations requiring discrimination, comparison, and judgment in adaptation.<sup>59</sup>

Sattler's classification system was as follows:

A. Language - the maturity of vocabulary, the extent of vocabulary, quality of vocabulary, and comprehension of verbal relations.

B. Memory - meaningful, non-meaningful and visual memory tests which reflect the auditory memory, ideational memory, and attention span.

C. Conceptual Thinking - primarily concerns abstract thinking. Closely associated with language ability.

D. Reasoning - verbal and non-verbal reasoning. The perception of logical relations, discrimination ability, and analysis and synthesis. Spatial reasoning is also measured.

E. Numerical Reasoning - arithmetic reasoning problems, numerical reasoning involves concentration and the ability to generalize from numerical data.

F. Visual-Motor - measures manual dexterity, eye-hand coordination, and perception of spatial relations. Also involves visual imagery and non-verbal reasoning.

---

<sup>59</sup>Robert Valett, Programming Learning Disabilities (Belmont, California: Fearon Publishers, 1969), p. 71.

G. Social Intelligence - measures social comprehension, social maturity, and social judgment.<sup>60</sup>

From these two classifications one saw similarities between Valett's General Comprehension and Sattler's Conceptual Thinking, between Valett's Vocabulary and Verbal Fluency and Sattler's Language, and between Valett's Judgment and Reasoning and Sattler's Conceptual Thinking as well as Social Judgment. "These two systems agree in classifying 75 percent of the total number of Stanford-Binet tests."<sup>61</sup> Classification systems could serve as provisional guides for grouping clusters in the scale in order to formulate hypotheses about the child's pattern of abilities and disabilities.

Wechsler Intelligence Scale for Children  
and the Wechsler Intelligence Scale  
for Children-Revised

The forces which had shaped the SBIS were part of the heritage of all tests. Therefore, it was essential that the material on the SBIS be understood in order to digest the knowledge concerning the Wechsler scales.

The Binet Scales having been introduced to the United States, discontent began to develop with the age-scale format. Yerkes, the leading spokesman, who with Bridges and Hardwick published the point scale in 1915,

. . . felt that tests on the SBIS were selected on the basis of proportions of successes and failures in selected age

---

<sup>60</sup>Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), p. 135.

<sup>61</sup>A. B. Silverstein, "Comparison of Two Item-Classification Schemes for the Stanford-Binet," Psychological Reports, XVII (1965), 964.

groups . . . this assumes that important forms of behavior appear at various points in development . . . in contrast, tests are selected for the point scale on the basis of their ability to measure various functions . . . that tests are selected according to percentage of passes and are grouped according to year level . . . where the point scale was more flexible . . . that the SBIS test uses an all or none scoring procedure for each test while the point scale uses a more-or-less scoring procedure.<sup>62</sup>

In designing the Wechsler-Bellevue Intelligence Scale, Form I, the forerunner to Form II and the WISC and Wechsler Preschool and Primary Scale of Intelligence (WPPSI), Wechsler, studying the standardized tests during the late 1930's, selected eleven different subtests to form the scale. Sources for the subtests included the Army Alpha for Information and Comprehension, the SBIS for Comprehension, Arithmetic, Digit Span, Similarities, and Vocabulary, the Healy Picture Completion Tests for Picture Completion, the Army Group Examinations for Picture Arrangement, and Kohs Block Design test for Block Design, and the Army Beta for Digit Symbol and Coding. Wechsler's search for subtests was guided by his conception of intelligence which supported the global nature of intelligence.

Wechsler included personality as a part of this large whole and designed the WISC to account for the total intelligence of the individual. No measure was taken to determine primary abilities or to put hierarchial priorities on any of the subtests. Wechsler's scales represented an index of general mental ability.

The WISC was a 1949 development as a downward extension of the Wechsler-Bellevue Intelligence Scale, and of Form II of the adult scales. In order for the WISC to be used with children, easier items are added

---

<sup>62</sup>Sattler, pp. 151-52.



to the low end of the subtests. The WISC was applicable to children between the ages of five years, zero months and fifteen years, eleven months. However, the WISC-R, which is the 1974 revision of the 1949 WISC, was applicable to children between the ages of six years, zero months to sixteen years, eleven months. However, there was an overlap of the WISC-R with the WPPSI (ages four to six and one-half) and the Wechsler Adult Intelligence Scale (WAIS) (ages sixteen to adult); either test could be used to evaluate children in this age range.

The WISC contained twelve subtests, six of which form the Verbal Scale and six of which form the Performance Scale. Two of the subtests, Digit Span and Mazes, were supplementary subtests. The WISC was standardized on 2200 white American boys and girls, considered to be representative of the 1940 U.S. census. There was an over-representation of children from the middle and upper socioeconomic levels. Wechsler appeared to reject the mental-age concept at first in that it did not represent an absolute level of mental capacity in regard to identical intelligence levels in different children. After the publication of the WISC, realizing the usefulness of mental-age equivalents, subsequent editions of the WISC manual provided a table of mental-age equivalents. Wechsler had considered the mental-age equivalents as guides to facilitate interpretation rather than a means of calculating intelligence quotients. The IQ was a deviation IQ with the mean being 100 and the standard deviation 15. Each subtest had a mean scaled score of 10 and a standard deviation of 3.

The distribution of the standardization sample of 2200 had provided the frame of reference for classifying IQ's. This is shown in Table 2, page 45.

Table 2  
Distribution of the 1949 Standardization Group

IQ	Classification
130 - up	Very superior
120 - 129	Superior
110 - 119	High average (bright)
90 - 109	Average
80 - 89	Low average (dull)
70 - 79	Borderline
69 - down	Mentally deficient (defective)

Bright, dull, and defective terms corresponded to bright normal, dull normal, and mental defective respectively, used in the WPPSI, WAIS, and 1949 WISC manual.<sup>63</sup>

#### Validity and Reliability

Validity and reliability studies provided support for the WISC as a valid and reliable instrument. The

Reliability coefficients for ages 7-1/2, 10-1/2, and 13-1/2 are .92 to .95 for the Full Scale, .88 to .96 for the Verbal Scale, and .86 to .90 for the Performance scale . . . subtests reliabilities range from a low of .59 for the Comprehension and Picture Completion subtests at the 7-1/2 age level to a high of .91 for the Vocabulary subtest at the 10-1/2 age level. The WISC and the Wechsler-Bellevue scales show a median correlation of .78 . . . the WISC and the WAIS scales show a median correlation of .84 . . . the WISC and the WPPSI scales show a median correlation of .81.<sup>64</sup>

The validity and reliability studies provide support for the WISC-R as a valid and reliable instrument. The

---

<sup>63</sup> David Wechsler, Manual for the Wechsler Intelligence Scale for Children-Revised (New York: Psychological Corporation, 1974), p. 26.

<sup>64</sup> Sattler, pp. 154, 155, 210.

Reliability coefficients are .95 to .96 for the Full Scale, .91 to .96 for the Verbal Scale and from .89 to .91 for the Performance Scale . . . subtests reliabilities range from a low of .57 to the Mazes subtests at the 16-1/2 age level to a high of .92 for the Vocabulary subtest at the 16-1/2 age level. The WISC-R and the WPPSI scales show a median correlation of .81 . . . the WISC-R and the WAIS scales show a median correlation of .84.<sup>65</sup>

### Classification System

Comparison among subtests was a type of classification that helped to describe the child's strengths and weaknesses by analyzing specific abilities when they were combined with other subtests. The hypotheses which one drew upon helped to facilitate the assessment process. In essence, the WISC and WISC-R had the same number and type of subtests, as well as being alike in description and classification. The main change with the WISC-R subtests as opposed to the WISC subtests was in content and administrative procedures. Wechsler's description and classification of the WISC and WISC-R subtests was as follows:

#### Verbal Scale

1. Information - measures the wealth of available information acquired as a result of native ability and early cultural experience. Memory is also an important aspect.
2. Comprehension - measures social judgment; the ability to use facts in a pertinent, meaningful, and emotionally relevant manner.
3. Arithmetic - measures one's reasoning ability plus numerical accuracy in mental arithmetic.
4. Similarities - measures verbal concept formation and logical thinking.

---

<sup>65</sup> Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), pp. 512, 514.

5. Vocabulary - measures a variety of functions, including language ability and fund of information.

6. Digit Span - determines the amount of one's attention and short term memory.

#### Performance Scale

1. Picture Completion - measures the ability to differentiate essential from non-essential details and requires concentration, visual organization, and visual memory.

2. Picture Arrangement - determines nonverbal reasoning ability as well as planning ability, the ability to comprehend and conceptualize a total situation.

3. Block Design - measures visual-motor coordination and perceptual organization.

4. Object Assembly - measures perceptual organization ability.

5. Coding - measures visual-motor coordination, speed of mental operation, and short-term memory.

6. Mazes - measures planning ability and perceptual organization.<sup>66</sup>

The Verbal Scale was highly structured and dependent on the individual's cumulative experience: it drew from what the child already knew. The Performance Scale was not highly structured and was more dependent on the individual's immediate problem-solving ability and required the student to meet new situations by applying past experiences and previously acquired skills to a new set of demands.

---

<sup>66</sup>Sattler, pp. 189-90.

### Revision of the WISC

The WISC-R was the 1974 revision of the 1949 WISC. It was similar to the WISC in that it contained twelve subtests which had the same names as those that appeared in the WISC.

The WISC-R was standardized on 2200 white and non-white American children who were selected as being representative on the basis of the 1970 U.S. census. The sample consisted of eleven different age groups ranging from six and one-half to sixteen and one-half years with 200 children in each group. Unlike the WISC, which did not include non-whites in the standardization group, the WISC-R included non-whites such as blacks, American Indians, Orientals, Puerto Ricans, and Mexican Americans in approximately the same proportions they represented in the U.S. population as shown by census data of 1970.

The WISC-R had a mean of 100 and a standard deviation of 15. The subtests had a mean scaled score of 10 and a standard deviation of 3.

The intercorrelations among the six Verbal Scale subtests had a median of .73 while those six Performance Scale subtests had a median of .48. The vocabulary subtests showed the highest correlation with the Full Scale (.74) as does the Block Design subtests (.68). The Digit Span subtest showed the lowest correlation with the Full Scale (.43) as does the Coding subtests (.38).

One of the difficulties with the WISC-R was that the range of the Full Scale IQ is 40 to 160; on the WISC it is 46 to 154. Therefore, unlike the Binet or the Slosson, children who had a mental age below six years, zero months or who were gifted may not be properly assessed. Part of this was due to the degree of points given which depends on the quality of the answer, and children received at least one scaled-score point even

though they had a raw score of 0. The only thing that could be done when this happened was in computing the IQ one could extrapolate the IQ on the basis of the sums of scaled scores or not compute a child's Full Scale IQ unless he obtained raw scores greater than 0 on at least three Verbal and three Performance Scale subtests. Some of the changes in the WISC-R are found in Table 3, pages 50-52.

Table 4, page 53, showed the relation of WISC-R's IQ's Scales Scores, and Scaled Scores Classifications to Deviations from the Mean and Percentile Ranks.

Percentile ranks for IQ's are shown in Table 5, page 54.

#### Slosson Intelligence Test

The SIT was an age-scale test that provided mental ages from .5 month to twenty-seven years. The items on the SIT were taken from the SBIS and the Gessell Institute of Child Development Behavior Inventory. Above age four all questions were presented verbally and spoken responses were required. It was a screening test which could be used by untrained and trained examiners. After several years of experimentation, only those items which produced favorable results were included. The SBIS was used as the criterion to validate the SIT. Even though the SIT "maintains the ratio IQ ( $MA/CA \times 100$ ), it has all the advantages of the deviation IQ."<sup>67</sup> This type of IQ had been used with the SBIS prior to the 1960 revision. The SIT was preferable to a group test and was a useful tool in selecting individuals for more comprehensive evaluation of

---

<sup>67</sup> Richard L. Slosson, Slosson Intelligence Test for Children and Adults (New York: Slosson Educational Publications, 1963), p. iv.

Table 3  
Changes in the WISC-R

Age range	WISC-R: 6-0-0 to 16-11-30 WISC: 5-0-0 to 15-11-30
Standardization	WISC-R: white and nonwhite Americans WISC: white Americans only
Number of subtests	WISC-R: 12 WISC: 12
Statistical data	WISC-R: Reliability coefficients, standard errors of measurement, and intercorrelations are available by yearly intervals from 6½ to 16½ and for the average of the 11 age groups. WISC: These data available only for three ages-- 7½, 10½, and 13½.
Content of subtests	(a) Same as WISC for Digit Span and Coding; (b) substantial changes in information, Similarities, Arithmetic, Vocabulary, Comprehension, Picture Completion, and Picture Arrangement; and (c) slight changes in Block Design, Object Assembly, and Mazes.
Subtest directions	Directions revised for administering all 12 subtests.
Scoring samples	Enlarged and clarified.
General administrative changes	(a) Verbal Scale subtests are alternated with Performance Scale subtests. (b) Examiner demonstrates the solution or provides the correct answer when the first item of a subtest is failed except for Digit Span and Coding. (c) Starting rules are changed for the Information, Arithmetic, and Vocabulary subtests, with different entry points for children aged 6 to 8, 8 to 10, 11 to 13, and 14 to 16. (d) Manual indicates that emotionally disturbed and other atypical children, as well as children below 8 years of age and older children suspected of mental deficiency, be started with the first item of each subtest.
Administrative changes for subtests	Information: Starting rules are changed. Specific responses that are incomplete to some items should be probed.

Table 3 (continued)

	<p>Similarities: Examples of 2-point responses are provided when child fails to give a 2-point response. All children started with item 1. Number of items increased from 16 to 17. Many more responses than on the WISC must be queried.</p>
	<p>Arithmetic: Starting rules are changed. Number of items increased from 16 to 18.</p>
	<p>Vocabulary: Starting rules are changed, all responses are scored 2, 1, or 0, and number of words is reduced from 40 to 32. Many more responses than on the WISC must be queried.</p>
	<p>Comprehension: A second idea is asked for when child gives only one idea on items requiring two ideas for full credit. Number of items increased from 14 to 17. Many more responses than on the WISC must be queried. Discontinuance rule changed to three consecutive failures.</p>
	<p>Digit Span: Both trials are always administered for each series of digits, and all series are scored 2, 1, or 0.</p>
	<p>Picture Completion: Maximum exposure is 20 seconds. Number of items increased from 20 to 26. Certain responses should be probed.</p>
	<p>Picture Arrangement: Child is encouraged to work quickly, and there are changes in the discontinuance rule (to three consecutive failures) and in scoring. Number of items increased from 11 to 12.</p>
	<p>Block Design: Blocks are two-colored, time limit on nine block design changed to 120 seconds, and allotment of bonus points and discontinuance procedures changed. Number of items increased from 10 to 11.</p>
	<p>Object Assembly: Scoring, including allotment of bonus points, and time limits are changed. Sample item is included. Correct solution is shown when the first item is failed.</p>
	<p>Coding: A separate booklet is used for the subtest. Minor changes in instructions.</p>



Table 3 (continued)

	Mazes: Scoring changed. Number of mazes increased from 8 to 9. Criteria for determining errors and instructions changed.
Scoring items failed below entry point	WISC-R: Manual states explicitly that credit is given to items failed below entry point items when entry point items are passed. WISC: No explicit statement appears in the manual about giving credit for items passed that appear below the entry point.
Scoring items passed above discontinuance point	WISC-R: Manual states explicitly that credit is not given when items are passed after the subtest should be discontinued. WISC: No explicit statement appears in the manual about not giving credit for items passed that appear above the discontinuance point.
Computation of IQ	(a) A Full Scale IQ should not be computed unless a raw score above 0 is obtained on at least three Verbal and three Performance Scale subtests. Similarly, a Verbal IQ is not computed unless raw scores above 0 are obtained on at least three Verbal Scale subtests. The same rule holds for computing the Performance IQ. (b) Digit Span and Mazes are not used to compute the IQ when the other 10 subtests have been administered. These two subtests were not included in the construction of the IQ table. (c) Days are used in determining the child's age.
Intelligence classification	"High Average" (110-119) is used in place of "Bright Normal," "Low Average" (80-89) in place of "Dull Normal," and "Mentally Deficient" (69 or below) in place of "Mental Defective."
Scaled scores not in "IQ Equivalents of Sums of Scaled Scores" table	Scales scores that do not appear in the table (either below or above the values in the table) should be reported as "under 40" or "over 160" for the Full Scale IQ and "under 45" or "over 155" for the Verbal or Performance Scale IQ.
Record booklet	Record booklet has discontinuance information, entry points shown by a black arrow with white numbers for age level, maximum points obtainable on each subtest, helpful hints for administering Similarities and Comprehension subtests, and a place for drawing a profile of scaled scores.*

\*Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), pp. 526-527.

Table 4

Relation of WISC-R IQ's Scales Scores and Scaled  
Scores Classification to Deviations from  
the Mean and Percentile Ranks

Verbal Performance, or Full Scale IQ	Scaled Score of Any Single Test	Number of SDs from the Mean	Percentile Rank	Classification
145	19	+3	99.9	Brilliant
140	18	+2-2/3	99.6	
135	17	+2-1/3	99	Very superior
130	16	+2	98	
125	15	+1-2/3	95	Superior
120	14	+1-1/3	91	
115	13	+1	84	Bright
110	12	+2/3	75	
105	11	+1/3	63	Normal
100	10	0 (Mean)	50	
95	9	-1/3	37	Dull
90	8	-2/3	25	
85	7	-1	16	Inferior
80	6	-1-1/3	9	
75	5	-1-2/3	5	Borderline deficient
70	4	-2	2	
65	3	-2-1/3	1	Deficient <sup>a</sup>
60	2	-2-2/3	0.4	
55	1	-3	0.1	

<sup>a</sup>David Wechsler, Manual for the Wechsler Intelligence Scale for Children-Revised (New York: Psychological Corporation, 1974), p. 25.

Table 5  
Percentile Ranks for IQ's

IQ	Percentile Rank		IQ	Percentile Rank	
	Stanford-Binet <sup>a</sup>	WISC-R, WISC, WPPSI, or WAIS		Stanford-Binet	WISC-R, WISC, WPPSI, or WAIS
135	99	98	99	48	48
134	98	98	98	45	45
133	98	98	97	43	43
132	98	98	96	40	40
131	97	98	95	38	37
130	97	97	94	35	35
129	96	97	93	33	32
128	96	96	92	31	30
127	95	96	91	29	28
126	95	95	90	27	26
125	94	95	89	25	24
124	93	94	88	23	22
123	93	93	87	21	20
122	92	92	86	19	18
121	90	91	85	17	16
120	89	90	84	16	15
119	88	89	83	15	13
118	87	88	82	13	12
117	85	87	81	12	11
116	84	85	80	11	10
115	83	84	79	10	9
114	81	82	78	8	8
113	79	80	77	7	7
112	77	78	76	7	6
111	75	76	75	6	5
110	73	74	74	5	5
109	71	72	73	5	4
108	69	70	72	4	4
107	67	68	71	4	3
106	65	65	70	3	3
105	62	63	69	3	2
104	60	60	68	2	2
103	57	57	67	2	2
102	55	55	66	2	2
101	52	52	65	1	1 <sup>b</sup>
100	50	50			

<sup>a</sup>Can be used with SIT IQ's.

<sup>b</sup>Jerome M. Sattler, Assessment of Children's Intelligence  
(Philadelphia: W. B. Saunders, 1974), p. 458.

mental ability. This test put heavy emphasis on verbal language skills as did the SBIS.

The population was drawn from the urban and rural areas in New York State. Referrals were from

. . . cooperative nursery schools, public, parochial and private schools, from junior and senior high schools . . . gifted as well as retarded classes . . . white, black, and some American Indians . . . others from city Youth Bureau, Home for Boys, professional groups, and county jails.<sup>68</sup>

The population standardization was approximately one thousand children.

The classification of IQ for the SIT is shown in Table 6.

Table 6

IQ Classification Chart

IQ	Classification	School Accomplishment and Placement
140 and up	Very Superior	Gifted classes, college, graduate work
120 - 139	Superior	Gifted classes, college, graduate work
110 - 119	Bright	High school and college
90 - 109	Average	High school, college is dubious
80 - 89	Dull	Slow learner classes
70 - 79	Borderline	Slow learner classes and classes for retarded
(Below 70 - Defective)		
50 - 69	Mild Retardation	Classes for retarded (Educable)
20 - 49	Moderate Retardation	Classes for retarded (Trainable)
0 - 19	Severe Retardation	School Exclusion for Trainable <sup>a</sup>

<sup>a</sup>Richard L. Slosson, Slosson Intelligence Test for Children and Adults (New York: Slosson Educational Publications, 1963), p. ii.

<sup>68</sup>Richard L. Slosson, Slosson Intelligence Test for Children and Adults (New York: Slosson Educational Publications, 1963), p. iv.

### Validity and Reliability

Validity and reliability studies provided support for the SIT as a useful screening instrument and as a device for retesting purposes. However, in no way was this test a substitute for the SBIS, WPPSI, WISC, or WISC-R.

Reliability coefficient suggests .97, .96, and .91. . . .  
 Validity coefficient suggests a range of .90 to .98 with the SBIS . . . .49 to .93 with the verbal scale of the WISC . . . .10 to .76 with the Performance scale of the WISC . . . . and from .50 to .84 with the full scale of the WISC (median correlation of .67) . . . these correlations may be spuriously high because the Slosson contains items that are essentially adaptations from the SBIS.<sup>69</sup>

### Classification System

There was no specific classification system for the SIT. However, since the SBIS had been used as a criterion to validate the SIT, it was suggested that Valett's system of classification for the Binet be used. This classification system was as follows:

- A. General Comprehension - the ability to conceptualize and integrate components into a meaningful total relationship.
- B. Visual-Motor Ability - the ability to manipulate material in problem-solving situations that usually require integration of visual and motor skills.
- C. Arithmetic Reasoning - the ability to make appropriate numerical associations and to deal with mental abstractions in problem-solving situations.
- D. Memory and Concentration - the ability to attend and retain. Requires motivation and attention and usually measures degree of retention of test items.

---

<sup>69</sup> Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), p. 246.

E. Vocabulary and Verbal Fluency - the ability to correctly use words in association with concrete or abstract material; the understanding of words and verbal concepts; the quality and quantity of verbal expression.

F. Judgment and Reasoning - the ability to comprehend and respond appropriately in specific situations requiring discrimination, comparison, and judgment in adaptation.<sup>70</sup>

### Standardized Achievement Tests

During the time that psychologists were developing intelligence and aptitude tests, traditional school examinations were undergoing technical improvements. In 1845 the Boston public schools substituted written examinations for oral responses of students by visiting examiners.

Horace Mann cited arguments concerning this innovation . . . which were used much later to justify the replacement of essay questions by objective multiple-choice items . . . these written examinations as noted by Mann put all students in a uniform situation, permitted a wider coverage of content, reduced the chance element in question choice, and eliminated the possibility of favoritism on the examiner's part.<sup>71</sup>

The first standardized tests for measuring the outcomes of school instruction began to appear as a result of the work of E. L. Thorndike. These tests utilized measurement principles that were the results of psychological laboratory experiments. The first set of achievement batteries was known to the public as the Stanford Achievement Test of

---

<sup>70</sup> Robert Valett, Programming Learning Disabilities (Belmont, California: Fearon Publishers, 1969), p. 71.

<sup>71</sup> Anne Anastasi, Psychological Testing (4th ed.; New York: Macmillan, 1976), p. 16.

1923. As these batteries showed increasing use in the schools, the shift of emphasis had been on the design of items to test the understanding and application of knowledge with regard to broad educational objectives. The decade of the 1930's witnessed the introduction of test-scoring machines for which the new objective tests could be readily adapted.

Later, the establishment of statewide, regional, and national testing programs developed. The best known of these programs was the College Entrance Examination Board (CEEB). In 1947 CEEB merged with the Carnegie Corporation and the American Council on Education to form Educational Testing Service (ETS). In the past ETS had assumed the responsibility for a growing number of testing programs on behalf of universities, professional schools, government agencies, and other institutions. The American College Testing Program was an important development; in 1959 it set out to screen applicants to colleges not included in the CEEB program and to select highly talented students for scholarship awards.

The main purpose of achievement tests was for educational use, but was not limited to this use. Many achievement tests were used in the selection of applicants for industrial and government jobs. However, as more psychologists were trained in psychometrics, the technical aspects of test construction in the achievement area more nearly resembled that of intelligence and aptitude tests.

The distinctive merit of standardized achievement testing is that it provides the teacher with an independent, objective yardstick that is less likely to reflect the teacher's own special biases in the coverage of a course . . . provides for evaluating the performance of his or

her class in the form of norms that are provided for such tests.<sup>72</sup>

A standardized achievement test had a further advantage of far more careful preparation and research than was ordinarily possible for individual teachers to provide for their own classroom examinations.

The role of measurement and evaluation in accountability established the fact that without some kind of accurate measurement there could be no valid evidence of the extent to which a program achieved its objectives. Without evaluation, no one was held accountable for his performance. With evaluation whoever was responsible for a given task was accountable for the results of his performance of that task. Standardized achievement tests were designed to measure objectives broader in scope than those sought in most performance objectives. In order to achieve valid accountability, the standardized achievement test was to measure the objectives of the educational program.

H. S. Adelman had implied that exceptional children are actually kindred in their educational achievement by traditional instruction because these children possess a unique array of learning characteristics that merit differential education.<sup>73</sup>

It is important that teachers view themselves as change agents of intelligence . . . psychoeducationally, they must know what a child does and does not know, how and under what conditions a child best learns and subsequently,<sup>74</sup> make some intelligent decision for educational remediation.

---

<sup>72</sup>Roger Lennon, Testing in the Secondary Schools, Test Service Notebook 20 (New York: Test Department, Harcourt, Brace, Jovanovich, 1968), p. 1.

<sup>73</sup>H. S. Adelman, "An Interactional View of Causality," Academic Therapy, VI (1970), 117-123.

<sup>74</sup>James S. Payne, "Psychoeducational Diagnosis," Mental Retardation Courses at the University of Virginia (Charlottesville: University of Virginia, 1974), pp. 1, 16.



Peabody Individual Achievement Test

In designing the PIAT between 1962 and 1969, Lloyd Dunn and Frederick Markwardt suggested that the purpose of the PIAT "is to provide a wide range screening measure of achievement in the areas of mathematics, reading, spelling, and general information."<sup>75</sup>

The PIAT was a power test, not a speed test. The five subtests were administered in a specific order. Mathematics was placed first because it did not require reading, writing, or oral response, and was a good rapport builder. The two reading subtests followed the mathematics subtest due to their crucial significance. Hoping the subject was maximally attentive, it was important to administer the Reading Recognition subtest before the Reading Comprehension subtest because word attack skills were usually a forerunner to getting meaning from sentences in print. The subject's performance on the Reading Recognition subtest determined whether or not he would be given the Reading Comprehension subtest. Spelling was the next subtest, followed by General Information. These two subtests, Spelling and General Information, were not included in the present study.

The Mathematics subtest was composed of items such as matching, discriminating, and recognizing numerals; it measured concepts in geometry and trigonometry as well as computational skills and arithmetic fundamentals.

---

<sup>75</sup>Lloyd M. Dunn and Frederick C. Markwardt, Jr., Manual for the Peabody Individual Achievement Test (Circle Pines, Minnesota: American Guidance Service, 1970), p. 1.

The Reading Recognition subtest was composed of items taken from vocabulary lists of basic reading series. The naming of letters, matching of letters, and the sounds of letters were included as part of this subtest.

The Reading Comprehension subtest was part of the Reading Recognition subtest, based on the theory that if one cannot achieve a certain level on the Reading Recognition subtest, he would not be able to pass any of the items on the Reading Comprehension subtest. The Reading Comprehension subtest was composed of items which the individual was required to explain based on the meaning of the sentence he had read.

The PIAT test was standardized

. . . from a national population of school children in the United States . . . the reason for this was to insure a cross-section of curriculum patterns . . . regular and special classrooms in public, private, and residential settings were included as well as three types of communities--urban, suburban, and rural . . . 15 to 25% of the standardized population were taken from special education facilities which yields a truncated distribution. . . . However, all students are taken from the mainstream of education . . . the total sample was 3,000 subjects spread equally over the 13 grade levels--kindergarten through 12th grade . . . there were 87 more girls than boys in the sample . . . this difference was not statistically significant . . . 84% were white, 11% were Negro, and 5% were other.<sup>76</sup>

The PIAT raw scores were converted into grade equivalents, age equivalents, percentile ranks and standard scores. The standard score was equivalent to a deviation IQ which has a mean of 100 and a standard deviation of 15. The PIAT also provided for the

. . . calculation of adjusted mental ages and the recording of the most current intelligence test data . . . as an index of the approximate level at which one could expect a student to achieve . . . when the intelligence test was

---

<sup>76</sup>Dunn and Markwardt, pp. 26-32.

given previous to the administration of the PIAT, it will be necessary to calculate the adjusted mental age by using the  $\frac{IQ}{100} \times C.A.$  (chronological age in months) = M.A. (Mental age in months) formula derived from the ratio IQ concept.<sup>77</sup>

### Validity and Reliability

Validity and reliability studies provided support for the PIAT as a useful screening instrument of achievement in mathematics, reading recognition, reading comprehension, spelling, and general information. The overall median reliability for this test was .89. The median reliability for mathematics was .74; for reading recognition was .89; for reading comprehension was .64; for spelling was .65 and general information was .76.

The overall median validity for this test was .68. The median validity for mathematics was .52; for reading recognition was .54; for reading comprehension was .66; for spelling was .40 and general information was .68. The PIAT was correlated with the Peabody Picture Vocabulary Test (PPVT) IQ scores which yielded a median correlation of .57. "Some 28 studies of relationships between the PPVT and other tests were reported with the median correlations with achievement tests in the .50s."<sup>78</sup>

### Classification System

The PIAT had no classification system such as the SBIS and the WISC-R uses. However, it did provide grade and age equivalents, percentile ranks, and standard scores.

---

<sup>77</sup>Dunn and Markwardt, p. 16.

<sup>78</sup>Dunn and Markwardt, p. 50.

### Significant Studies

Some of the most significant studies concerning validity were as follows:

1. Kennedy, Van de Riet, and White, in 1963, found a correlation of .69 between the SBIS (Form L-M) MA and the California Achievement Test scores. This was a sample of 1800 Negro elementary school children between the ages five and sixteen who were living in the southeastern United States. The SBIS (Form L-M) was found to correlate .67 with the Metropolitan Achievement Test in a sample of disadvantaged children, 80 percent of whom were Negroes between the ages of nine and eleven years. Significant correlations of .38 to .61 had been found between the WISC and Wide Range Achievement Test (WRAT) in samples of seven-year-old white as well as Negro children as reported by Henderson, Butler, and Gofferey in 1969.<sup>79</sup>
2. Studies comparing the SBIS and the Wechsler Adult Intelligence Scale (WAIS) in samples of mentally retarded individuals were high, ranging from .74 to .90 with a median correlation of .75. Studies comparing samples of normal college freshmen showed that the WAIS and SBIS correlations range from .40 to .83 with a median correlation of .77.<sup>80</sup>
3. The bulk of studies had used the SBIS as the criterion for evaluating the comparative validity of the WISC. Forty-seven studies indicated the SBIS and WISC were highly comparable by showing correlation coefficients ranging from .44 to .92 for the WISC (VIQ) and SBIS; .30 to

---

<sup>79</sup> Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), pp. 43-44.

<sup>80</sup> Sattler, p. 125.

.86 for the WISC (PIQ) and SBIS; .43 to .94 for the WISC (FSIQ) and SBIS as well as a median correlation of .80.<sup>81</sup>

4. The WISC and Wechsler-Bellevue studies indicated a range of .41 to .82 for the Performance scale; .54 to .90 for the Verbal Scale; .70 to .89 for the Full Scales with a median correlation of .78. The WAIS, which was the Wechsler-Bellevue replacement, suggests correlations of .66 to .96 for the Verbal Scale; .51 to .92 for the Performance Scale; .70 to .95 for the Full Scale with a median correlation of .84 when compared to the WISC.<sup>82</sup>

5. Other studies used the WISC as the criterion against which other tests were validated. Correlations between the WISC (FSIQ) and the Columbia Test of Mental Maturity, Draw-a-Man, and Progressive Matrices suggested a range of .49 to .74 with a median correlation of .62 with the Columbia Test of Mental Maturity; .04 to .59 with a median correlation of .36 with the Draw-a-Man, and .27 to .91 with a median correlation of .51 with the Progressive Matrices.<sup>83</sup>

6. Several studies using the WISC revealed good validity relative to scores on academic achievement. A variety of children were used, and correlations revealed a range of .14 to .81 for the WISC (FSIQ) with a median correlation of .01.<sup>84</sup>

7. The Stanford-Binet had been most popular in serving as the criterion for the WPPSI. Thirteen studies indicated that correlations

---

<sup>81</sup>Sattler, p. 155.

<sup>82</sup>Sattler, p. 155.

<sup>83</sup>Sattler, p. 155.

<sup>84</sup>W. M. Littell, "The Wechsler Intelligence Scale for Children: Review of a Decade of Research," Psychological Bulletin, LVII (1960), 132-56.

between the WPPSI and SBIS range from .33 to .92 with a median correlation of .81 for the verbal scale; .33 to .88 with a median correlation of .67 for the Performance Scale, and from .44 to .92 with a median correlation of .82 for the full scale.<sup>85</sup>

8. The WPPSI and the WISC had correlations of .57 to .91 for the Verbal Scale; .43 to .82 for the Performance Scale; .65 to .90 for the Full Scale (median correlation of .81). The WPPSI and other tests such as the Progressive Matrices and the Primary Mental Abilities Test had correlations from a low of .30 to a high of .82 respectively. The median correlation was .64.<sup>86</sup>

9. Correlations between the WPPSI and two reading tests, the Gilmore Oral Reading Paragraphs Test and the Stanford Achievement Test for Reading in the prediction of reading achievement in first grade suggested a .57 and .61 correlation for the Verbal Scale respectively; .58 and .63 correlations for the Performance Scale respectively; .62 and .68 correlation for the Full Scale respectively; the two highest subtest correlations were arithmetic and geometric design yielding a correlation of .63.<sup>87</sup>

10. The SBIS and the WISC were useful in diagnostic processes concerning special children. Considerable evidence from these tests helped in the diagnostic analysis of childhood schizophrenia and early infantile autism, in the assessment of organic brain damage, in the assessment of and predictive power for mental retardates. The WISC and SBIS did not show any systematic patterns that could distinguish

---

<sup>85</sup> Sattler, p. 209.

<sup>86</sup> Sattler, p. 210.

<sup>87</sup> Sattler, pp. 232-33.

emotionally disturbed children from normal children or from children with other forms of psychopathology. The WISC and SBIS showed predictive usefulness with learning disability children as to cluster patterns and significant discrepancies between their estimated intellectual potential and their actual level of performance. With physical disabilities or physical diseases some effect on intelligence occurred in blindness and deafness, but was more adverse in cerebral palsy, symptomatic epilepsy, and muscular dystrophy.<sup>88</sup>

11. In a survey of the preferences of school psychologists in California, Weise in 1960 found

. . . that the SBIS was preferred in testing the gifted and mentally retarded for grades K-2 . . . starting with the 3rd grade the WISC was preferred in testing for mental retardation . . . in testing for giftedness in grades 3-6, there was equal preference for each of the two tests . . . after grade 6 the WISC was preferred for the gifted . . . the WISC was also preferred to the SBIS in evaluating problems concerned with differential diagnosis (e.g., learning problems, emotional problems, or neurological problems). . . .<sup>89</sup>

12. Correlations between the WISC-R and WPPSI suggested .80 for the Verbal Scales; .80 for the Performance Scales; .82 for the Full Scales; the .82 correlation was almost identical with the median correlation of .81 that had been found in studies comparing the WISC and WPPSI.<sup>90</sup>

13. Correlations between the WISC-R and WAIS suggested .96 for the Verbal scales; .83 for the Performance Scales; .95 for the Full Scales; the .95 correlation was higher than the median correlation of .84 that had been found comparing the WISC and WAIS.<sup>91</sup>

---

<sup>88</sup> Sattler, pp. 283, 301, 320, 344, 345.

<sup>89</sup> Sattler, p. 413.

<sup>90</sup> Sattler, p. 514.

<sup>91</sup> Sattler, p. 514.

14. Four groups of children ages six, nine and one-half, twelve and one-half, and sixteen and one-half, who were included in the WISC-R standardization sample were also administered the SBIS. The correlations between the WISC-R and SBIS are as follows:

<u>Scales</u>	<u>Ages</u>				<u>Average</u>
	<u>6</u>	<u>9-1/2</u>	<u>12-1/2</u>	<u>16-1/2</u>	
Verbal Scale	.77	.64	.66	.73	.71
Performance Scale	.74	.57	.51	.51	.60
Full Scale	.82	.69	.63	.74	.73

These results were similar to those found in studies concerning the WISC and SBIS. The WISC-R Vocabulary subtest correlated more highly with the SBIS (.69) than any of the other subtests; the Coding subtest had the lowest correlation (.26) of any of the subtests. The SBIS norms were based on the 1972 standardization.<sup>92</sup>

15. Correlations between the SBIS and the Reading, Arithmetic, and Language part of the California Achievement Test were .68, .64, and .78 respectively were obtained with a large sample of black elementary school children.<sup>93</sup>

16. Correlations with the WAIS (VIQ) and college or engineering school grades, had been between .40 to .50. Another study involving the WAIS and SBIS with unselected adolescent or adult groups as well as mental retardates clustered around a correlation of .80. In other studies the WAIS (PIQ) correlated .70 with Raven's Progressive Matrices

---

<sup>92</sup>Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), p. 515.

<sup>93</sup>Anne Anastasi, Psychological Testing (4th ed.; New York: Macmillan, 1976), p. 244.



and .72 between the WAIS (PIQ) and the Minnesota Form Board Test in a group of sixteen-year-old boys and girls.<sup>94</sup>

17. Correlations between the Differential Aptitude Tests given early in high school at the tenth grade level to be used to predict status near the end of grade twelve on the College Entrance Examinations Board Scholastic Aptitude Tests for 169 boys and girls on the verbal portion and 199 boys and 119 girls on the numerical portion showed correlations of .79 for the SAT-V actual and predicted scores, and .85 for the SAT-N actual and predicted scores.<sup>95</sup>

18. The effect of reading disability upon intelligence test performance was shown in a study by Neville in 1965, when he compared individual WISC scores and Lorge-Thorndike verbal test scores for good, mediocre, and poor readers in a fifth grade class. Reading groups were defined by scores on the reading section of the Metropolitan Achievement Test at the end of the fourth grade. Results indicated a thirty-one IQ point difference between good readers and poor readers; an eighteen IQ point difference between good readers and mediocre readers; and a thirteen IQ point difference between mediocre readers and poor readers for the Lorge-Thorndike verbal test. The WISC (VIQ) results indicated a twenty-one IQ point difference between good readers and poor readers; a fourteen IQ point difference between good readers and mediocre readers; a seven IQ point difference between mediocre readers and poor readers.<sup>96</sup>

---

<sup>94</sup>Anastasi, Psychological Testing, p. 253.

<sup>95</sup>H. Seashore, "Tenth Grade Tests as Predictors of Twelfth Grade Scholarship and College Entrance Status," Journal of Counseling Psychology, I (1954), 106-15.

<sup>96</sup>Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education (3d ed.; New York: John Wiley and Sons, 1969), pp. 317-18.

19. Wide Range Achievement Test (WRAT) scores and California Mental Maturity IQ's had been correlated, and for the WRAT Reading subtest and the WRAT Arithmetic subtest correlations coefficients were .81 and .84 respectively. Correlations between the WRAT Reading and WAIS (VIQ) were .84; between the WRAT Arithmetic and WAIS (VIQ) were .76; between the WRAT Reading and WAIS (PIQ) were .60; between the WRAT Arithmetic and WAIS (PIQ) was .67; between the WRAT Reading and WAIS (FSIQ) was .76; between the WRAT Arithmetic and WAIS (FSIQ) was .77. Correlations between the WISC and WRAT were as follows:

	WISC (average)		
	(VIQ)	(PIQ)	(FSIQ)
WRAT Reading	.71	.52	.68
WRAT Arithmetic	.69	.60	.74

It was interesting to note that for boys and girls the Information and Vocabulary subtests on the WISC had the highest correlation with the WRAT Reading, and the Arithmetic and Information subtests on the WISC had the highest correlation with the WRAT Arithmetic.<sup>97</sup>

20. Correlations between intelligence and the Key Math Test showed a .59 correlation for a group of forty-five educable mentally retarded adolescents; with a group of twenty-eight normal fifth graders were .38 for the full scale Iowa arithmetic score and .69 with the reasoning measure on the Iowa Test. Both were significant at the .05 level.<sup>98</sup>

---

<sup>97</sup>J. F. Jastak and S. R. Jastak, Manual of Instructions: Wide Range Achievement Test (Wilmington, Delaware: Guidance Associates, 1965), p. 18.

<sup>98</sup>Austin J. Connolly, William Nachtman, and E. Milo Pritchett, Manual for the Key Math Diagnostic Arithmetic Test (Circle Pines, Minnesota: American Guidance Service, 1971), p. 30.

21. The Full-Range Picture Vocabulary Test (FRPV) suggested .76 with the SBIS for sixty male mental defectives; .82 with the WISC for ninety child reading cases; .84 with SBIS vocabulary for eighty Negro children; .67 and .69 with the SBIS vocabulary for 360 school children.<sup>99</sup>

22. The relationship between the California Mental Maturity IQ's and the WRAT suggested a correlation of .81 for Reading and .84 for Arithmetic . . . WISC and WAIS scores in relationship to the WRAT scores showed correlations of .85 for Reading and .75 for Arithmetic using the WISC (FSIQ) for 200 boys ages nine to eleven; .82 for Reading and .81 for Arithmetic using the WISC (FSIQ) for 200 girls ages nine to eleven; the WAIS (FSIQ) showed a .84 correlation on the WRAT Reading and .83 correlation on the WRAT Arithmetic for 200 males ages eighteen to twenty-four; the WAIS (FSIQ) showed a .87 correlation on the WRAT Reading and .83 correlation on the WRAT arithmetic for 200 females ages eighteen to twenty-four.<sup>100</sup>

23. The most comprehensive study of school achievement was completed by the U.S. Public Health Service, based on a national sample of school age children from the first to twelfth grade. The Stanford and Metropolitan Achievement Tests were compared with the WRAT Reading and Arithmetic subtests. The mean correlation for all grades and all tests was .70.<sup>101</sup>

---

<sup>99</sup>W. Sloan and G. J. Bensberg, "An Exploratory Study of the Full Range Picture Vocabulary Test with Mental Defectives," American Journal of Mental Deficiency, LVIII (1954), 481-85; L. M. Smith and A. R. Fillmore, "The Ammons FRPV Test and the WISC for Remedial Reading Cases," Journal of Consulting Psychology, XVIII (1954), 332; N. W. Coppinger and R. B. Ammons, "The Full-Range Picture Vocabulary Test: VIII, A Normative Study of Negro Children," Journal of Clinical Psychology, VII (1952), 136-40; R. B. Ammons, P. R. Arnold, and R. S. Herrmann, "The Full Range Picture Vocabulary Test: IV, Results for a White School Population," Journal of Clinical Psychology, VI (1950), 164-69.

<sup>100</sup>J. F. Jastak and S. R. Jastak, Manual of Instructions: Wide Range Achievement Test (Wilmington, Delaware: Guidance Associates, 1976), pp. 51-52.

<sup>101</sup>Jastak and Jastak, Manual of Instructions, 1976, pp. 53-57.

24. Sitlington in 1970 compared the PIAT with the WRAT using educable mentally retarded adolescents achieving at the third grade level. In general the PIAT subtests

. . . correlated more highly with their WRAT counterparts than with any of the other measures; PIAT Mathematics vs WRAT Arithmetic, .58; PIAT Reading Recognition vs WRAT Reading, .95; PIAT Spelling vs WRAT Spelling, .85; PIAT Reading Comprehension and General Information have no WRAT counterparts. PIAT Reading Comprehension correlated quite highly with WRAT Reading showing a .90 correlation.<sup>102</sup>

25. Studies using the PPVT with varied populations involving handicapped children showed median correlations of .71 and .83 with SBIS mental age scores on the 1937 and 1960 series respectively; other median correlations showed .61 with the WISC (FSIQ); .67 with the WISC (VIQ); .39 with the WISC (PIQ); .79 with the WAIS (FSIQ); .84 with the WAIS (VIQ); .62 with the WAIS (PIQ). The PPVT correlated well with scholastic achievement tests such as the Sequential Tests of Educational Progress (STEP), the California Achievement Tests (CAT), the Wide Range Achievement Test (WRAT), the Stanford Achievement Test (SAT), and the Metropolitan Achievement Tests (MAT). The median correlations value was .50.<sup>103</sup>

26. M. S. Swanson and A. Jacobson found a correlation of .64 between the WISC (VIQ) and the SIT on sixty-four suburban second graders referred for learning problems. A correlation of .10 was found on the WISC (PIQ) which suggested that the SIT was essentially a measure of verbal intelligence.<sup>104</sup>

---

<sup>102</sup>Lloyd M. Dunn and Frederick C. Markwardt, Jr., Manual for the Peabody Individual Achievement Test (Circle Pines, Minnesota: American Guidance Service, 1970), p. 51.

<sup>103</sup>Lloyd M. Dunn, Manual for the Peabody Picture Vocabulary Test (Circle Pines, Minnesota: American Guidance Service, 1965), pp. 33-41.

<sup>104</sup>M. S. Swanson and A. Jacobson, "Evaluation of the Slosson Intelligence Test for Screening Children with Learning Disabilities," Journal of Learning Disabilities, III (1976), 318-20.

27. R. H. Pate and W. R. Nichols found a correlation of .84 between the SIT IQ and the WISC (FSIQ) on one hundred students referred for special class evaluation. These students were suggested for mentally retardation class placement. It was concluded that the SIT was a useful screening device.<sup>105</sup>

28. Kaufman and Ivanoff found a correlation of .93 between the SIT and the WAIS when used with a rehabilitation population.<sup>106</sup>

29. Armstrong and Reynolds, using a sample of 198 elementary school students having been referred for special class placement, found correlations between the SIT and WISC (VIQ) to be .93; between the SIT and WISC (PIQ) to be .75; between the SIT and WISC (FSIQ) to be .90. Evidence showed that the SIT, WISC (VIQ) and WISC (FSIQ) are good estimates of each other. Using the SBIS, Armstrong, Mooney, and Jensen used a homogeneous population of 147 special class students ranging in age from eight to fourteen and in IQ from 45 to 88 with a mean of 72. The correlation between the SIT and the SBIS was .85. In another study by Armstrong and Jensen involving 490 students age six to fourteen, who were enrolled in ten public school systems, showed a correlation of .93 between the SIT and the SBIS. These findings suggested that the SIT could be used as a valid screening and retesting substitute for the SBIS.<sup>107</sup>

---

<sup>105</sup> R. H. Pate, Jr. and W. R. Nichols, "The Slosson Intelligence Test Used as a Screening Device: A Validity Study," American Psychological Association, Catalog of Abstracts, Experimental Publication System, 1970, No. 9, p. 10.

<sup>106</sup> H. Kaufman and J. Ivanoff, "The Slosson Intelligence Test as a Screening Instrument with a Rehabilitation Population," Exceptional Children, XXXV (1969), 745.

<sup>107</sup> Slosson, attached abstracts, p. 27.

30. A study by Wilson and Spangler, using the PIAT with a population sample of eighty-three children and adolescents having been referred because of learning difficulties, attempted to assess the value of the PIAT as an effective measure and screening device of educational achievement. The CA range was five years, three months to eighteen years, three months and each individual was administered one or more standardized test of intelligence as part of a large array of evaluations. The decision as to whether to administer a WISC, SBIS, or PPVT was a decision that was made by the psychologist. The correlations between the PIAT grade levels and IQ scores were as follows:

PIAT and WISC = .58

PIAT and SBIS = .49

PIAT and PPVT = .45

The results of this study suggested that the PIAT could be used for both elementary and adolescent children who presented problems in the areas of mental retardation, learning disabled, emotionally disturbed, and sensory problems. The moderate positive correlations were consistent with other reports. Sitlington in 1970 found a correlation of .58 to .95 between the PIAT and WRAT for a group of forty-six EMR adolescents functioning at the third grade level. Soethe in 1972 found the following correlations using the PIAT and the WISC and WRAT for a population of forty children labeled normal, reading disabled, and mentally retarded. The correlations were as follows:

PIAT + WRAT for mentally retarded = .44 - .87 range

PIAT + WRAT for reading disabled = .37 - .92 range

PIAT + WISC (FSIQ) for mentally retarded = .22

PIAT + WISC (FSIQ) for normals = .80

It was suggested that the PIAT + WISC + WRAT did aid in differential diagnosis of learning problems.<sup>108</sup>

31. Validity studies showed correlations between Slosson IQ and Stanford-Binet IQ to be in the range from .90 (age four) to .90 (ages six and seven); between Slosson and Wechsler Full Scale IQ's range from .54 to .93; between Slosson and Stanford-Binet from .76 to .90; between Slosson and Wechsler Adult Intelligence Scale Full Scale IQ's for rehabilitation clients was .93; .96 between the Slosson and verbal IQ on the Wechsler Adult Intelligence Scale, and .70 between the Slosson and Performance IQ on the Wechsler Adult Intelligence scale. Studies with a sample of fifty-six pupils with reading problems showed a correlation of .64 and .42 between the Slosson and the verbal and performance IQ on the Wechsler Intelligence Scale for Children; a study compared the Slosson with both the Wechsler Intelligence Scale for Children and the Stanford-Binet with retarded children which showed correlations of .54, .85, and .20 between the Slosson and Wechsler Intelligence Scale for Children for the Full Scale IQ, Verbal IQ, and Performance IQ respectively. The Stanford-Binet correlations are .76 and .81 with mental ages, and .79 with IQ's.<sup>109</sup>

32. Studies investigating the validity of standardized intelligence tests with populations of ethnic minority groups usually report validity coefficients which are similar to those obtained with white populations . . . studies show correlations of .69 between the Stanford-Binet mental age and

---

<sup>108</sup> John D. Wilson and Paul F. Spangler, "The Peabody Individual Achievement Test as a Clinical Tool," Journal of Learning Disabilities, VII (1974), 60-63.

<sup>109</sup> Oscar Buros, ed., The Seventh Mental Measurements Yearbook, Volumes I and II (Highland Park, New Jersey: Gryphon Press, 1972), pp. 764-67.

California Achievement Test scores . . . .64 with grades in academic areas and the Stanford-Binet . . . .67 between the Stanford-Binet (Form L-M) and the Metropolitan Achievement Test; .57 with reading achievement on the same test . . . the Peabody Picture Vocabulary Test have been found to be related significantly to reading achievement scores . . . as well as be a valid predictor for first grade children . . . significant correlations between .38 to .61 have been found between the Wechsler Intelligence Scale for Children and the Wide Range Achievement Test in samples of seven year old white as well as Negro children.<sup>110</sup>

### Summary

This chapter provides a review of related literature relative to an introduction as well as administrative decision-making, the nature of intelligence and psychological testing, the measurement of intelligence, intelligence tests, standardized achievement tests, significant studies, and a summary.

---

<sup>110</sup>Jerome M. Sattler, Assessment of Children's Intelligence (Philadelphia: W. B. Saunders, 1974), pp. 43-44.



## Chapter 3

### METHODOLOGY

#### Introduction

A major purpose of this research involved prediction. Three intelligence tests and an achievement test were used in this study as aids in facilitating some outcomes relative to educational decisions.

This chapter includes the methodology employed in the study relative to the population, sample and criteria used in drawing the sample. Tests are identified and described as are the criteria established for selecting the three mental maturity tests and one achievement test. The method applied and summary are also included.

#### Population

This school division served approximately 10,000 students with a teaching staff of about 675 professionals resulting in a pupil-staff ratio of approximately 15-1. In addition an administrative staff of approximately 100 provided a variety of services from the central office.

The population included three levels of school organization: There were fifteen elementary schools that served grades kindergarten through five ranging in size from 100 to 650 enrollment; four middle schools that served grades six through eight ranging in size from 400 to 800 enrollment, and three secondary schools that served grades nine through twelve ranging in size from 1,000 to 2,000 students. The third secondary school was a Vocational Technical Education Center.

The school division provided a wide variety of sequential and comprehensive special education programs in the area of special services for approximately 1,200 handicapped children and youth ages two to twenty-one and grades pre-school through grade twelve. This was approximately 12 percent of its student population.

Included in this school division's continuum of services were programs for children and youth identified as mentally retarded, physically handicapped, emotionally disturbed, learning disabled, speech impaired, hearing impaired, multiple handicapped, and visually impaired. A full range of supportive services was available, as well as pre-school programs for handicapped children.

#### The Sample

The study sample was limited to no more than the first 200 students in grades one through five and no more than the first 100 students in grades six through twelve referred to Special Services by their teachers, guidance counselor or principal.

The criterion used in drawing the sample from the population was that all of the potential 300 students referred were presumed to have some type of handicapped condition.

This school division had a system of referral. There were eighteen steps relative to the identification, evaluation, confirmation, and placement of special education students (see Appendix A). There were fifteen steps relative to the reevaluation of special education students (see Appendix B).

### Tests Identified and Described

The standardized instruments used in this study included three mental maturity tests and one achievement test. The three predictors in the study were the Stanford-Binet Intelligence Scale (SBIS), the Wechsler Intelligence Scale for Children-Revised (WISC-R), and the Slosson Intelligence Test (SIT).

The criterion used in the study was the Peabody Individual Achievement Test (PIAT).

The criteria used for selecting the three mental maturity tests and one achievement test were as follows:

1. Reference data as to title, author, publisher, type of test, cost, and test time needed to administer the tests.

2. Data of manual, names of tests and subtests, types of profiles given as well as test items included in the test, purpose of test, qualifications needed to administer, score, and interpret the test.

3. Type of validity and reliability reported, method and adequacy of item sampling, statistical procedure used to describe validity, sampling procedure for determining test reliability, type and strength of reliability and validity coefficient, mean and standard deviations.

4. Subtest inter-correlations, time limits, types of norms provided, characteristics of norm group, method of sampling norm groups, and reviewer's comments.

A description of each test follows:

1. Stanford-Binet Intelligence Scale (SBIS). The SBIS was an individual test of intelligence. The third revision (1960) form L-M consisted of one form used to measure intelligence of individuals from age two to adult. It was also used with handicapped persons. Formal training was needed before one was allowed to administer such a test. It was newly normed in 1972. If the examiner compared how the individual stood in relation to his current age group, the 1972 norms were to be used. The time required to administer this test varied from thirty to ninety minutes. The SBIS maintained the deviation IQ (mean 100 and standard deviation 16) and the mental age (MA). To score the SBIS a basal age was established by taking that level at which all tests were passed which just precedes the level where the first failure occurred. Upon reaching a ceiling level, which was the level at which all tests were failed, the examiner calculated the mental age (MA) by taking the basal age and added the earned credits to it. There were abbreviated tests starred which were used if the examiner wanted to use a short form test and there were alternative tests that were used as a substitute when a test had been spoiled. However, no alternative test could be substituted for a test which had been failed.

There was a specific classification system for the SBIS, based on Valett's and Sattler's work. The six areas were general comprehension (conceptual thinking according to Sattler), visual-motor ability, arithmetic reasoning, memory and concentration, vocabulary and verbal fluency (language according to Sattler), and judgment and reasoning (conceptual thinking and social judgment according to Sattler).

2. Wechsler Intelligence Scale for Children-Revised (WISC-R). The WISC-R was an individual test of intelligence. It measured verbal

and performance (non-verbal) tasks, and it was a revision (1974) and restandardization of the 1949 WISC.

The WISC-R was used with individuals whose age range was six years to sixteen years, eleven months. It was also used with handicapped persons. Formal training was needed before one was allowed to administer such a test. The WISC-R contained twelve subtests. Two of the subtests were supplementary and were not included in the verbal, performance, and full scale IQ when given. The supplementary test in the verbal area was digit span, and in the performance area was mazes. It was recommended that the ten subtests be given alternately between verbal and performance areas. The subtests were given in the following order: information, picture completion, similarities, picture arrangement, arithmetic, block design, vocabulary, object assembly, comprehension, and coding. The time required to administer this test varied from fifty to seventy-five minutes. The WISC-R maintained the deviation IQ (mean 100 and standard deviation 15) and used no mental age (MA). However, test age equivalents were given for subtests only.

To score the WISC-R different directions were given for each subtest as to scoring, where to start and stop according to the age of the child, and the quality of one's answer. Time was an additional factor to consider on the performance items. Raw scores were converted to scale scores and from a table the verbal IQ, performance IQ, and full scale IQ were determined. The subtest scale scores had a mean of ten and a standard deviation of three.

There was a specific classification system for the WISC-R. It was based on each subtest, and additional interpretation was possible by grouping certain subtests as clusters.

3. Slosson Intelligence Test (SIT). The SIT was a short individual test of intelligence, used as an individual screening instrument with individuals from ages two weeks to adult. It was also used with handicapped persons. No formal training was needed. The items were adopted from the SBIS. The time required to administer this test varied from ten to thirty minutes. The SIT maintained the ratio IQ ( $MA/CA \times 100$ ) and mental age (MA) even though it had the advantages of a deviation IQ. To score the SIT a basal age was established by taking the highest level of successful passes before the first error after ten in a row passes. Upon reaching a ceiling level, which was the highest level where ten in a row were missed, the examiner calculated the mental age (MA) by taking the basal age and adding earned credits to it.

There was no specific classification system for the SIT. Since the SBIS had been used as a criterion to validate the SIT, it was suggested that Valett's system of classification be used.

4. Peabody Individual Achievement Test (PIAT). The PIAT was a wide range individual screening test of achievement in the areas of mathematics, reading recognition and reading comprehension, spelling, and general information. This test was also used as a diagnostic instrument with individuals whose age ranges were from kindergarten through adult. The PIAT was a power test not a speed test, and it required thirty to forty minutes to administer. The PIAT was especially useful for the handicapped person because it required no reading or writing--just pointing. The subtests were given in order as follows: mathematics, reading recognition, reading comprehension, spelling, and general information. The items within each subtest were arranged in ascending order of difficulty. To score the PIAT a basal age was established by

correctly answering five consecutive responses prior to the first error. The ceiling level was reached when there existed five errors in any seven consecutive responses. The critical range extended from the basal item to the ceiling item, and the raw score was the number of errors over the critical range subtracted from the number of the ceiling item.

#### Method Applied

The entire sample was tested using the SIT, SBIS, WISC-R, and PIAT by professionally qualified examiners and certified psychologists. The three mental maturity tests were administered within a three week period from the time of referral. The time schedule limits for the administration of the three selected mental maturity measures was from October 1977 through March, 1978. The PIAT was administered during May of 1978. The level of significance established for this study was .05. The statistical test applied was the Pearson Product-Moment Correlation. The resulting  $\gamma$  statistic was the validity coefficient  $\gamma_{xy}$  which was the correlation between predictor and criterion. The higher the validity coefficient, the greater was the correlation between the two variables.

There were four types of derived scores. Grade and age equivalent scores were the developmental types, and percentile ranks and standard scores were the deviation types. The standard score was equivalent to a deviation IQ which had a mean of 100 and a standard deviation of 15. The PIAT also provided for the calculation of an adjusted mental age by using the  $MA = \frac{IQ}{100} \times CA$  when the intelligence test had been given prior to the administration of the PIAT.

The PIAT had no specific classification system. However, any of the derived scores were very useful when comparing the achievement to IQ and mental age (MA).

### Summary

The organization of the collected data relative to analysis was done in the following manner: The name of the student referred was listed; the time of referral was coded; the grade at the time of referral was listed; the specific referral handicap was recorded; the individual's chronological age at the time of referral and at the time the criterion test was given were recorded, as was the appropriate group checked depending on whether the student was at the elementary or secondary level. IQ's and MA's were recorded for the three mental maturity tests; grade equivalents and MA's were recorded for the achievement test in mathematics and reading. An adjusted (MA) was recorded for all predictor tests. A sample copy of this collected data worksheet was in Appendix C.

This chapter provided a review of methodology relative to an introduction as well as population, the sample and criteria used in drawing the sample, tests identified and described as well as the criteria established for selecting the three mental maturity tests and one achievement test, method employed, and a summary.



## Chapter 4

### ANALYSIS OF DATA

#### Introduction

Implicit in the concept of criterion-related validity was that tests were useful and increased the accuracy of the decision-making process. The type of criterion-related validity used in this study was predictive validity which purported to predict future performance (the criterion) on the basis of present performance (the predictor). The test used for the criterion was the PIAT. The test used for the three predictors were the SBIS, WISC-R, and the SIT. The criterion test measured mathematics and reading achievement. The three predictors measured mental maturity.

The categories of information presented were the introduction, the sample, primary hypothesis and the sub-hypotheses.

#### The Sample

The sample in this study was composed of two groups. The first group represented 140 elementary students in grades one through five. The second group represented 91 secondary students in grades six through twelve. The total sample population was 231 students who were referred to Special Services because each was presumed to have a handicapped condition.

The percentages of the sample population handicapped categories for Group I, Group II, and Groups I and II are presented in Table 7.

Table 7  
Percentages of the Sample Population  
Handicapped Categories

Handicapped Category	Group I (N = 140)		Group II (N = 91)		Groups I and II (N = 231)	
	No.	Percent	No.	Percent	No.	Percent
Learning Disabled	112	80.0	52	57.1	164	70.9
Educable Mentally Retarded	12	8.6	15	16.5	27	11.7
Emotionally Disturbed	7	5.0	3	3.3	10	4.3
Re-evaluation Educable Mentally Retarded	7	5.0	14	15.4	21	9.1
Re-evaluation Trainable Mentally Retarded	2	1.4	6	6.6	8	3.4
Re-evaluation Learning Disability	-	-	1	1.1	1	0.6
Totals	140	100.0	91	100.0	231	100.0

Inspection of this table revealed that the handicaps not included in the sample population handicapped categories were the physically handicapped, the speech impaired, the hearing impaired, the multiple handicapped, the visually impaired, and the re-evaluation of the learning disabled for Group I.

The handicapped category which revealed the highest percentage of the sample population was the learning disabled category for Group I, and Group II, and Groups I and II.

The handicapped category which revealed the lowest percentage of the sample population was the re-evaluation of the trainable mentally retarded category for Group I, and the re-evaluation of the learning disabled for Group II and Groups I and II.

Another interesting insight related to the sample population was the numbers and percentages by schools as to the time of referral. The three mental maturity tests were administered within a three week period from the time of referral. The time schedule limits for the administration of the three selected mental maturity measures was from October 1977 through March 1978. The achievement test was administered during May of 1978.

The numbers and percentages by schools as to the time of referral are presented in Table 8.

Table 8  
Numbers and Percentages by Schools of the Sample  
Population as to Time of Referral

Schools	Oct.-Dec. 15 1977		Jan.-Feb. 15 1978		Feb. 16-Mar. 31 1978		Totals	
	No.	Percent	No.	Percent	No.	Percent	No.	Percent*
Group I								
Broadus Wood	12	5.2	8	3.5	11	4.8	31	13.4
Greer	17	7.4	14	6.1	18	7.8	49	21.2
Hollymend	12	5.2	10	4.3	13	5.6	35	15.2
Scottsville	10	4.3	9	3.9	6	2.6	25	10.8
Group II								
Albemarle	4	1.7	3	1.3	10	4.3	17	7.4
Jouett	9	3.9	8	3.5	14	6.1	31	13.4
Walton	14	6.1	16	6.9	13	5.6	43	18.6
Totals	78	33.8	68	29.4	85	36.8	231	100.0

\* Group I was 60.6 percent of total sample  
Group II was 39.4 percent of total sample

Inspection of Table 8 revealed that 60.6 percent of the total sample was represented by Group I, and 39.4 percent of the total sample was represented by Group II.

The largest number of referrals occurred during February 16 to March 13, 1978 for the total sample. The smallest number of referrals occurred during January to February 15, 1978 for the total sample.

In Group I, Greer had the largest percentage of referrals while Scottsville had the smallest percentage. In Group II, Walton had the largest percentage of referrals while Albemarle had the smallest percentage.

A descriptive analysis of the sample for this study are presented in Table 9.

Table 9

Mean Age (Criterion) for Both Groups, Raw Score Means and Standard Deviations of the Predictors for Both Groups, and Adjusted Mental Age Means and Standard Deviations of the Predictors and Criterion for Both Groups

	Group I (N = 140)				Group II (N = 91)			
	Raw Score		Adjusted Mental Age		Raw Score		Adjusted Mental Age	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Age (criterion)	9-1				14-2			
IQ's (predictors)								
SIT	96.70	16.22	8-7	2-0	85.33	20.11	12-0	2-6
SBIS	92.21	16.14	8-3	2-0	82.45	18.18	11-7	2-5
WISC-RV	94.13	16.37	8-5	2-0	85.79	18.05	12-2	2-5
WISC-RP	93.76	17.46	8-5	2-3	88.23	17.34	12-5	2-7
WISC-RFS	93.36	17.01	8-5	2-1	85.99	18.00	12-2	2-6
Achievement (criterion)								
Mathematics (PIAT)			8-3	2-0			10-3	2-9
Reading (PIAT)			7-8	1-4			9-7	2-3

Inspection of Table 9 revealed that the mean age (criterion) for Group I was nine years, one month (fourth grader), and for Group II was fourteen year, two months (ninth grader).

Raw score means and standard deviations of the predictors for Group I reflected similar scores to those standardized means and standard deviations for the SBIS and WISC-R. The SIT used a ratio IQ instead of a deviation IQ.

Raw score means of the predictors for Group II reflected a standard deviation below the standardized means for the SBIS and WISC-R. However, raw score standard deviations of the predictors for Group II reflected similar scores to those standardized standard deviations for the SBIS and WISC-R. The SIT used a ratio IQ instead of a deviation IQ.

The adjusted mental age means and standard deviations of the predictors and criterion provided interesting insights. The adjusted mental age mean was equivalent to a third grader for the predictors of Group I and was equivalent to a seventh grader for all predictors of Group II except the SBIS predictor, which was equivalent to a sixth grader. All standard deviations for the adjusted mental ages of both Group I and II possessed similar variability.

The adjusted mental age mean of the mathematics criterion for Group I was equivalent to a third grader, and for Group II that of a fifth grader.

The adjusted mental age mean of the reading criterion for Group I was equivalent to a second grader, and for Group II that of a fourth grader.

Standard deviations for the adjusted mental ages for both groups reflected similar variability with reading in Group I having the smallest variability.

The sample population in Group I had a mean achievement level in mathematics equivalent to their mean potential, but when compared to their mean expected grade norm for their age (criterion), it was found to be one year below the expected mean grade placement level.

The sample population in Group I had a mean achievement level in reading one year below their mean potential, but when compared to their mean expected grade norm for their age (criterion), it was found to be two years below the expected mean grade placement level.

The sample population in Group II had a mean achievement level in mathematics two years below their mean potential except for the SBIS predictor, which revealed one year below the mean potential, but when compared to their mean expected grade norm for their age (criterion), it was found to be four years below the expected mean grade placement level.

The sample population in Group II had a mean achievement level in reading three years below their mean potential except for the SBIS predictor, which revealed two years below the mean potential; but when compared to their mean expected grade norm for their age (criterion), it was found to be five years below the expected mean grade placement level.

### Primary Hypothesis

A computer program was written for this data analysis by analysts at East Tennessee State University's Office of Computer Services.

The following primary hypothesis was tested at the .05 confidence level:

$H_0$ : There was no significant difference in the predictive ability between the SIT, SBIS, and the WISC-R, when compared to academic achievement in mathematics and reading as measured on the PIAT.

Table 10 contained the Pearson Product Moment Correlation Coefficients for Groups I and II between PIAT raw scores on the mathematics and reading subtests and the SIT, SBIS, and WISC-R (V), (P), (FS) IQ raw scores. The null hypothesis was rejected. Predictive ability between the SIT, SBIS, and the WISC-R were different due to differences in the validity coefficients. The higher the validity coefficients, the greater was the correlation and predictive ability. All tests were significant at the .05 level.

Table 10

Pearson Product Moment Correlation Coefficients Between  
PIAT Raw Scores on the Mathematics and Reading  
Subtests and the SIT, SBIS, and WISC-R(V),  
(P), (FS) IQ Raw Scores

Group I = Grades 1-5 (N = 140)\*

Group II = Grades 6-12 (N = 91)\*

Tests	. $\gamma$ for Mathematics	. $\gamma$ for Reading
SIT - PIAT		
Group I	.39	.34
Group II	.70	.70
SBIS - PIAT		
Group I	.38	.39
Group II	.75**	.77**
WISC-R(V) - PIAT		
Group I	.37	.38
Group II	.75**	.76
WISC-R(P) - PIAT		
Group I	.48**	.46**
Group II	.65	.60
WISC-R(FS) - PIAT		
Group I	.47	.46**
Group II	.75**	.73

\* All  $\gamma$ 's significant at .05 level

\*\* Test yielding highest validity coefficient

Inspection of Table 10 revealed that the WISC-R(P) was the best predictor of mathematics ability for Group I; the WISC-R(P) and the WISC-R(FS) were the best predictors of reading ability for Group I; the SBIS, WISC-R(V), and WISC-R(FS) were the best predictors of mathematics ability for Group II; the SBIS was the best predictor of reading ability for Group II.

The correlation coefficients were generally higher for Group II. The WISC-R(P) held the most substantial relationship in both mathematics and reading for Group I and the SBIS held the most substantial relationship in both mathematics and reading for Group II.

#### Sub-hypothesis

The following sub-hypothesis was tested at the .05 confidence level:

$H_0$ : There were no significant differences between mental ages as derived from SIT, SBIS, and the WISC-R, and the MA as derived from grade placement on mathematics and reading of the PIAT.

1. There was no significant difference between the SIT MA and the Reading MA.

2. . . . between the SIT MA and the mathematics MA.

3. . . . between the SBIS MA and the reading MA.

4. . . . between the SBIS MA and the mathematics MA.

5. . . . between the WISC-R(V) MA and the reading MA.

6. . . . between the WISC-R(V) MA and the mathematics MA.

7. . . . between the WISC-R(P) MA and the reading MA.

8. . . . between the WISC-R(P) MA and the mathematics MA.

9. . . . between the WISC-R(FS) MA and the reading MA.

10. . . . between the WISC-R(FS) MA and the mathematics MA.



Table 11 contains the Pearson Product-Moment Correlation Coefficients between PIAT mental age scores on the mathematics and reading subtests and the SIT, SBIS, and WISC-R(V), (P), and (FS) mental age scores. The null hypothesis was rejected. Mental ages as derived from SIT, SBIS, and the WISC-R, and the MA as derived from grade placement on mathematics and reading of the PIAT were different due to differences in the validity coefficients. The higher the validity coefficients, the greater was the correlation and predictive ability. All tests were significant at the .05 level.

Table 11

Pearson Product Moment Correlation Coefficients Between  
PIAT Mental Age Scores on the Mathematics and  
Reading Subtests and the SIT, SBIS, and  
WISC-R(V), (P), and (FS) Mental Age  
Scores

Group I = Grades 1-5 (N = 140)\*

Group II = Grades 6-12 (N = 91)\*

Tests	. $\gamma$ for Mathematics	. $\gamma$ for Reading
SIT - PIAT		
Group I	.77**	.73
Group II	.75	.79
SBIS - PIAT		
Group I	.74	.75
Group II	.74	.82**
WISC-R(V) - PIAT		
Group I	.75	.76
Group II	.76**	.79
WISC-R(P) - PIAT		
Group I	.74	.73
Group II	.51	.55
WISC-R(FS) - PIAT		
Group I	.77**	.77**
Group II	.69	.73

\* All  $\gamma$ 's significant at .05 level

\*\* Test yielding highest validity coefficient

Inspection of Table 11 revealed that the SIT and WISC-R(FS) were the best predictors of mathematics ability for Group I; the WISC-R(FS) was the best predictor of reading ability for Group I. The WISC-R(V) was the best predictor of mathematics ability for Group II; the SBIS was the best predictor of reading ability for Group II.

The correlation coefficients were generally high for both Groups I and II. The WISC-R(FS) held the most substantial relationship in both mathematics and reading for Group I and the WISC-R(V) and SBIS held the most substantial relationship in mathematics and reading for Group II respectively.

Table 12 contained the Pearson Product-Moment Correlation Coefficients for Group I and Group II among the intelligence test raw scores.

Inspection of Table 12 revealed that the SIT had the highest correlation with the SBIS for both Group I and Group II; the SBIS had the highest correlation with WISC-R(V) for both Groups I and II; the WISC-R(V) had the highest correlation with the WISC-R(FS) for Group I and with the SBIS for Group II; the WISC-R(P) had the highest correlation with the WISC-R(FS) for both Groups I and II; the WISC-R(FS) had the highest correlation with the WISC-R(V) for both Groups I and II.

The correlation coefficients were generally higher for Group II. Group I had the highest correlation with the WISC-R on four out of five intelligence tests, and Group II had the highest correlation with the WISC-R on three out of five intelligence tests. Group I had the highest correlation with the SBIS on one out of five intelligence tests, and Group II had the highest correlation with the SBIS on two out of five intelligence tests. Therefore, 70 percent of the time, regardless of

of which group, the WISC-R had the highest correlation with the other intelligence tests, and 30 percent of the time, regardless of which group, the SBIS had the highest correlation with the other intelligence tests.

Table 12

Pearson Product Moment Correlation Coefficients for  
Groups I and II Among the Intelligence Tests  
Raw Scores (N = 231)

Group I = Grades 1-5 (N = 140)\*  
Group II = Grades 6-12 (N = 91)\*

Test and Group	SIT	SBIS	WISC-R V	WISC-R P	WISC-R FS
SIT					
Group I		.83**	.80	.64	.78
Group II		.94**	.92	.74	.89
SBIS					
Group I	.83		.91**	.70	.87
Group II	.94		.97**	.77	.93
WISC-R(V)					
Group I	.80	.91		.71	.93**
Group II	.92	.97**		.77	.95
WISC-R(P)					
Group I	.64	.70	.71		.92**
Group II	.74	.77	.77		.93**
WISC-R(FS)					
Group I	.78	.87	.93**	.92	
Group II	.89	.93	.95**	.93	

\* All  $\gamma$ 's significant at .05 level

\*\* Test yielding highest validity coefficient

### Summary

This chapter provided a review of the analysis relative to an introduction, the sample, the primary hypothesis, the sub-hypothesis, and summary.

## Chapter 5

### SUMMARY

#### Introduction

The problem of the study was to determine by comparison which of the three selected mental maturity measures was the best predictor of achievement in mathematics and reading as measured by the Peabody Individual Achievement Test.

The primary hypothesis stated there was no significant difference in the predictive ability between the SIT, SBIS, and the WISC-R when compared to academic achievement in mathematics and reading as measured on the PIAT.

The sub-hypothesis stated there were no significant differences between mental ages, as derived from SIT, SBIS, and the WISC-R, and the MA as derived from grade placement on mathematics and reading of the PIAT.

There was a comprehensive search of the literature; criteria were established for the purpose of selecting the three mental maturity tests and the one achievement test to be used in the study; criteria used in drawing the sample from the population were that all of the potential students referred were presumed to have some type of handicapped condition.

The entire sample was tested using the SIT, SBIS, WISC-R and PIAT by professionally qualified examiners and certified psychologists. The three mental maturity tests were the three predictors, and the achievement test was the criterion in the areas of mathematics and reading.

The primary and sub-hypotheses were tested in the null format, and the level of significance established for this study was .05. The statistical test used in this study was the Pearson Product-Moment Correlation. The resulting  $r$  statistic was the validity coefficient  $r_{xy}$  which was the correlation between predictor and criterion. The higher the validity coefficient, the greater was the correlation between the two variables.

An analysis of the research was conducted, and categories of information relative to this study follow in sequence as to findings, conclusions, and recommendations.

### Findings

The study sample was composed of two groups. The first group represented 140 elementary students in grades one through five. The second group represented 91 secondary students in grades six through twelve. The total sample population was 231 students. It was found in Groups I and II that the learning disability handicapped category had the highest referral percentage showing it was 80 percent of the sample for Group I and 57 percent of the sample for Group II, giving an overall 70 percent of the entire sample referred. Other handicaps included in the sample were the educable mentally retarded, emotionally disturbed, and the trainable mentally retarded. Handicaps not included in the sample were the physically handicapped, the speech impaired, the hearing impaired, the multiple handicapped, and the visually impaired. The handicapped categories that had the lowest referral percentage were the re-evaluation of the trainable mentally retarded for Group I, and the re-evaluation of the learning disabled for Group II and Groups I and II.

The proportion of the total sample represented by Group I was 60.6 percent, and 39.4 percent of the total sample was represented by Group II.

The largest number of referrals (85) occurred during February 16 to March 31, 1978 for the total sample. The smallest number of referrals (68) occurred during January to February 15, 1978 for the total sample. In Group I, Greer had the largest percentage of referrals (21.2 percent) while Scottsville had the smallest percentage of referrals (10.8 percent). In Group II, Walton had the largest percentage of referrals (18.6 percent) while Albemarle had the smallest percentage of referrals (7.4 percent).

The mean age (criterion) for Group I was nine years, one month (fourth grade), and for Group II was fourteen years, two months (ninth grader).

Raw score means and standard deviations of the predictors for Group I reflected similar scores to those standardized means (100) and standard deviations for the SBIS (16) and WISC-R (15). The SIT used a ratio IQ instead of a deviation IQ. Group II reflected a standard deviation below the standardized means (100) for the SBIS and WISC-R, but reflected similar scores to those standardized standard deviations for the SBIS (16) and the WISC-R (15). The SIT used a ratio IQ instead of a deviation IQ.

The adjusted mental age mean for Group I was equivalent to a third grader for all predictors, and in Group II was equivalent to a seventh grader for all predictors except the SBIS, which showed a sixth grade equivalent. All standard deviations for the adjusted mental ages of both Group I and Group II possessed similar variability.

The adjusted mental age mean of the mathematics criterion for Group I was equivalent to a third grader and for Group II that of a fifth grader.

The adjusted mental age mean of the reading criterion for Group I was equivalent to a second grader, and for Group II that of a fourth grader.

Standard deviations for the adjusted mental ages for both groups reflected similar variability, with reading in Group I having the smallest variability.

The sample population for Group I had a mean achievement level in mathematics equivalent to their mean potential, but it was one year below the expected mean grade placement level. The mean achievement level for Group II in mathematics was two years below their mean potential, except for the SBIS predictor, which revealed one year below the mean potential but it was four years below the expected mean grade placement level.

The sample population for Group I had a mean achievement level in reading one year below their mean potential, but it was two years below the expected mean grade placement level. The mean achievement level for Group II in reading was three years below their mean potential, except for the SBIS predictor, which revealed two years below the mean potential, but it was five years below the expected mean grade placement level.

The null hypothesis was rejected for the primary hypothesis. Predictive ability between the SIT, SBIS, and the WISC-R was different due to differences in the validity coefficients. The higher the validity coefficients the greater was the correlation and predictive ability.



Pearson Product Moment Correlation coefficients between PIAT raw scores on the mathematics and reading subtests and the SIT, SBIS, and WISC-R (V), (P), (FS) IQ raw scores revealed that the WISC-R(P) was the best predictor of mathematics ability for Group I; the SBIS, WISC-R(V) and WISC-R(FS) were the best predictors of mathematics ability for Group II. The WISC-R(P) and the WISC-R(FS) were the best predictors of reading ability for Group I. The SBIS was the best predictor of reading ability for Group II. Generally, correlation coefficients were higher for Group II than Group I. The WISC-R(P) held the most substantial relationship in both mathematics and reading for Group I, and the SBIS held the most substantial relationship in both mathematics and reading for Group II.

The null hypothesis was rejected for the sub-hypotheses. Mental ages as derived from SIT, SBIS, and the WISC-R, and the MA as derived from grade placement on mathematics and reading of the PIAT were different due to differences in the validity coefficients. The higher the validity coefficients the greater was the correlation and predictive ability.

Pearson Product-Moment Correlation coefficients between PIAT mental age scores on the mathematics and reading subtests and the SIT, SBIS, and WISC-R(V), (P), and (FS) mental age scores revealed that the SIT and WISC-R(FS) were the best predictors of mathematics ability for Group I; the WISC-R(V) was the best predictor of mathematics ability for Group II. The WISC-R(FS) was the best predictor of reading ability for Group I. The SBIS was the best predictor of reading ability for Group II. Generally, correlation coefficients were high for both groups. The WISC-R(FS) held the most substantial relationship in both

mathematics and reading for Group I, and the WISC-R(V) and SBIS held the most substantial relationship in mathematics and reading for Group II respectively.

### Conclusions

The handicapped category which revealed the highest percentage of the sample population was the learning disabled category for Groups I and II. Eighty percent of the sample for Group I was learning disabled and 57 percent of the sample for Group II was learning disabled. This resulted in 70 percent of the entire sample population being learning disabled.

The criterion used in drawing the sample from the population was that all of the potential students referred were presumed to have some type of handicapped condition. This study revealed that the population was actually more of a learning disability population and was mostly characteristic of that type of handicap. Evidence of this was revealed in Table 9, where there was an educationally significant discrepancy between their estimated intellectual potential and actual level of performance. There existed a greater discrepancy between their actual level of performance and their expected grade placement level, these specific learning disabilities appeared to be related to basic disorders in the learning process.

The schools which had the largest and the smallest percentage of referrals in both Group I and II also had the largest school population from which to draw referrals. The exception to this was Albemarle in Group II which had the largest school population, but it had the smallest percentage of referrals in Group II. Some causes for this were

student dropouts, students moving out of the system, students too old for the administration of the WISC-R(V), (P), and (FS), students who had learned to compensate for their difficulties and were not referred, and students whose problems were felt to be beyond remediation due to age and the complexity of the processes involved.

The sample population's mean achievement level for Group I in mathematics was compared to their mean expected grade norm for their age (criterion), and found to be one year below the expected mean grade placement level; when the sample population's mean achievement level in reading was one year below their mean potential and two years below the expected mean grade placement level, this was indicative of a learning disability population. This group's mean IQ met the single most generally agreed-upon requisite for learning disabled classification which was that of an average (normal) level of intellectual functioning. Eighty percent of the sample for Group I was classified in the learning disabled handicapped category.

The sample population's mean achievement level for Group II in mathematics was compared to their mean potential, and it was found to be two years below their mean potential, except for the SBIS which revealed one year below the mean potential. However, compared to their expected mean grade placement level, it was found to be four years below grade norm for their age (criterion).

The mean achievement level in reading was found to be three years below their mean potential except for the SBIS, which revealed two years below the mean potential, and when compared to their expected mean grade placement level, it was found to be five years below grade norm for their age (criterion).

This group's mean IQ did not meet the single most generally agreed upon requisite for learning disabled classification which was that of an average (normal) level of intellectual functioning. The range of their mean IQ reflected low average or often referred to as dull normal or slow learner categories. Although 57 percent of the sample for Group II was referred as learning disabled, 38.5 percent of the sample for Group II had scores that fell within the moderate to borderline retardation range which tended to lower the overall mean IQ and achievement levels. This was evident by the two years in mathematics and three years in reading differentiation compared to their mean potential which was indicative of a learning disability and borderline retardation population. However, it was evident by the four years in mathematics and five years in reading differentiation compared to their expected mean grade placement level that this was indicative of mild and moderate retardation population.

Another factor that existed was that the older the child was the more the predictors were associated with conceptual, abstract, sequential, comprehension, reasoning, and spatial type of tasks as well as psycholinguistic processes.

It was interesting to evidence the Pearson Product-Moment Correlation coefficients for Group I and II among the intelligence tests' raw scores, which showed a higher validity coefficient for Group II than Group I.

It was also evident that the WISC-R as a predictor had the highest validity coefficient and the most substantial relationship 80 percent of the time for Group I and 60 percent of the time for Group II. The SBIS as a predictor had the highest validity coefficient and the most substantial relationship 20 percent of the time for Group I and 40 percent of the time for Group II.

The percentage differences in favor of the WISC-R at the Group I levels as opposed to the Group II levels was attributed to the fact that the WISC-R was more global in nature tapping many more areas of verbal and non-verbal learning, whereas the SBIS taps very little non-verbal learning but emphasized verbal, abstract, conceptual, and comprehension types of tasks.

Pearson Product-Moment Correlation coefficients between PIAT raw scores on the mathematics and reading subtests and the SIT, SBIS, and WISC-R(V), (P), (FS) IQ raw scores showed that Group II had higher validity coefficients than Group I. However, Pearson Product Moment Correlation coefficients between PIAT mental age scores on the mathematics and reading subtests and the SIT, SBIS, and WISC-R(V), (P), and (FS) mental age scores showed that Group I and Group II validity coefficients were similar.

Reasons for the differences between the groups with raw scores compared to the lack of differences between the groups with mental age scores were the way the different tests were structured, scored, the younger the child the less he had to know to acquire points. The advantage of a mental age score over an IQ score was that it derived the actual performance level of the individual in terms of the chronological age (CA) at which that score was an average score.

Comparisons of the groups with IQ achievement raw scores with those groups with IQ achievement mental age scores showed that 75 percent of the time the WISC-R was the best predictor and held the most substantial relationship for IQ achievement raw scores and IQ achievement mental age scores. The only group where this was not the result was Group II on the reading sub-tests for both IQ achievement raw scores and IQ achievement

mental age scores. The SBIS was the best predictor and held the most substantial relationship 25 percent of the time. The only other predictor was the SIT for Group I on the mathematics subtests in the comparison of IQ achievement mental age scores.

Comparing the mean age (criterion) for both groups, raw score means and standard deviations of the predictors for both groups, and adjusted mental age means and standard deviations of the predictors and criterion for both groups as well as in the comparisons between the IQ achievement raw scores and IQ achievement MA scores showed that the WISC-R(P) was greater than the WISC-R(V) for the majority of the cases. This was expected of students who were classified learning disabled due to difficulties in the area of mathematics and reading. Often subtest profiles for learning disabled children had pronounced peaks and valleys with the verbal comprehension and conceptual categories associated with the WISC-R(V) component and reading having the lowest scores, as did the sequential and distractibility categories, which were associated with the WISC-R(V) component and mathematics. The sequential category was also reflective of short term storage and retrieval of sequences of visual and auditory stimuli which were skills critically associated with reading ability. Most of the foundation for mathematics achievement was not limited to visuospatial skills, but also involved language concepts.

It was not surprising to find these results in relation to the WISC-R, since 70 percent of the entire sample population was learning disabled.

The SBIS, which had the most substantial relationship 25 percent of the time in conjunction with both Group II's on the reading

subtests for IQ achievement raw scores and IQ achievement mental age scores, showed to be a highly verbal type of test where reading competence was more dependent on the development and mastery of perceptual skills in the lower grades and early years of a child's school life, and in the higher grades and later years of a child's school life, reading competence was more dependent on conceptual factors such as comprehension and reasoning. The SBIS offers very few non-verbal types of tasks compared to the WISC-R(P) types of tasks. At the same time, the SBIS required more in-depth type of verbal conceptualizations than the WISC-R(V) types of tasks.

It was not surprising to find that overall the WISC-R was the best predictor of mathematics and reading, because the WISC-R provided for a more global and gestalt appraisal of those verbal and non-verbal types of tasks necessary in the processing of information relative to reception, association, and expression. These tasks are imperative in the acquisition of mathematics and reading skills.

Criterion-related validity studies implied that standardized tests were used as part of the decision-making process. Therefore, while the criterion used in this study showed excellent correlation with its predictors, it was crucial to this study that the type of standardized achievement test chosen be not only appropriate for the type of sample it was to be administered to, but that it be also sensitive to curriculum differences. Too often student achievement in a particular curriculum was not reflected by achievement test scores. All standardized achievement measures did not representatively sample different curricula. Such biases had to be acknowledged and considered any time a standardized, norm referenced achievement test was used for decision making.

### Recommendations

The following recommendations suggested additional studies of a predictive validity nature which would be helpful in the decision-making process:

1. It was suggested that a study determining the different time intervals the predictors and criterion were administered would show any significant difference as to validity coefficient correlations between the mental maturity tests and the achievement test in mathematics and reading.

2. It was suggested that a study be conducted to determine which predictor at which grade level (one through twelve) best predicts achievement in mathematics and reading.

3. It was suggested that a study similar to the present one be considered with the sample being a stratified random sampling in that equal samples and handicapped conditions from each grade one through twelve were essentially equal in size and handicapped classifications. It was recommended that forty-five be the number in each stratified sample with each of the nine handicaps having five members of each handicap at each grade level. This would suggest 225 in Group I and 315 in Group II giving a total stratified random sampling of 540.

4. It was suggested that a study determining a predictability equation be attempted by making multiple correlations where the three predictors would be used as one predictor in determining achievement in mathematics and reading. However, additional variables would be added if desirable such as an aptitude test, a mathematical and/or reading achievement test taken at some grade level prior to the present one, or one's semester grade point average.



5. It was suggested that a study using the regression equation be implemented where a prediction of an individual's performance on a criterion variable was made from the individual's performance on a predictor's variable. In other words, given an IQ score, what would one's achievement level be in mathematics and reading?

6. It was suggested that a longitudinal study be made using the PIAT to determine environmental factors on learning and the aging process on the acquisition and retention of various learned skills and knowledge.

7. It was suggested that a study using the comprehension, spatial, conceptual, and sequential subtest clusters of the WISC-R be correlated separately with mathematics and reading achievement using three IQ subgroups divided as to high, average, and low. The Illinois Test of Psycholinguistic Ability (ITPA) and the Detroit Test of Learning Aptitude (DTLA) would be used as process predictors with the WISC-R subtest clusters also used as predictors, or would be used as additional criteria with mathematics and reading achievement, which would be correlated with the WISC-R subtest clusters used as predictors.

These seven suggested studies would add to the findings of the present study in helping to increase the accuracy of the decision-making process, in the implementation of evaluation strategies of greater depth and quality relating to the student's needs as a process of prediction, and in sustaining accountability through educational goals and outcomes. It was predicted from this study that these three areas could be enhanced by the recommended additional studies because all correlation validity coefficients were significant at the .001 level; this was interpreted to mean that only once in 1000 times would the observed correlation be attributable to chance factors.

## BIBLIOGRAPHY

## BIBLIOGRAPHY

- Adelman, H. S. "An Interactional View of Causality," Academic Therapy, VI (1970), 117-123.
- Ammons, R. B., P. R. Arnold, and R. S. Herrmann. "The Full Range Picture Vocabulary Test: IV, Results for a White School Population," Journal of Clinical Psychology, VI (1950), 164-69.
- Anastasi, Anne. Individual Differences. New York: John Wiley and Sons, 1965.
- \_\_\_\_\_. Psychological Testing. 4th ed. New York: Macmillan, 1976.
- \_\_\_\_\_. "Psychology, Psychologists, and Psychological Testing," American Psychologists, XXII (1967), 297-306.
- Barnard, Chester I. Functions of the Executive. Cambridge, Massachusetts: Harvard University Press, 1938.
- Brown, Frederick. Principles of Educational and Psychological Testing. Hinsdale, Illinois: The Dryden Press, 1970.
- Buros, Oscar, ed. The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1965.
- \_\_\_\_\_, ed. The Seventh Mental Measurements Yearbook. Volumes I and II. Highland Park, New Jersey: Gryphon Press, 1972.
- Connolly, Austin J., William Nachtman, and E. Milo Pritchett. Manual for the Key Math Diagnostic Arithmetic Test. Circle Pines, Minnesota: American Guidance Service, 1971.
- Coppinger, N. W., and R. B. Ammons. "The Full-Range Picture Vocabulary Test: VIII, A Normative Study of Negro Children," Journal of Clinical Psychology, VII (1952), 136-40.
- Cronbach, Lee J. Essentials of Psychological Testing. 3d ed. New York: Harper and Row, 1970.
- Division of Special Services. Guidance Handbook for Virginia Schools, XLVII, No. 11. Richmond, Virginia: State Department of Education, June 1965.
- Dunn, Lloyd M. Manual for the Peabody Picture Vocabulary Test. Circle Pines, Minnesota: American Guidance Service, 1965.
- \_\_\_\_\_, and Frederick C. Markwardt, Jr. Manual for the Peabody Individual Achievement Test. Circle Pines, Minnesota: American Guidance Service, 1970.

Ebel, R. "Declining Scores: A Conservative Explanation," Phi Delta Kappan, IV (1976), 309.

Education Section, Division of Special Education. Administrative Requirements and Guidelines for Special Education Programs. Richmond, Virginia: State Department of Education, May 1972.

Galton, Francis. Inquiries into Human Faculty and Its Development. London: Macmillan, 1883.

Griffiths, Daniel E. Administrative Theory. Englewood Cliffs, New Jersey: Prentice-Hall, 1959.

Jastak, J. F., and S. R. Jastak. Manual of Instructions: Wide Range Achievement Test. Wilmington, Delaware: Guidance Associates, 1965.

\_\_\_\_\_. Manual of Instructions: Wide Range Achievement Test. Wilmington, Delaware: Guidance Associates, 1976.

Johnson, C. H., Jr. "Court, Craft, and Competence: A Reexamination of Teacher Evaluation Procedures," Phi Delta Kappan, IX (1976), 606.

Kaufman, H., and J. Ivanoff. "The Slosson Intelligence Test as a Screening Instrument with a Rehabilitation Population," Exceptional Children, XXXV (1969), 745.

Knezevich, Stephen J. Administration of Public Education. 3d ed. New York: Harper and Row, 1975.

Knudsen, C. Evaluation and Improvement of Teaching. New York: Doubleday, Doran, 1932.

Lennon, Roger. Testing in the Secondary Schools. Test Service Notebook 20. New York: Test Department, Harcourt, Brace, Jovanovich, 1968.

Lessinger, Leon M. "Accountability: Present Forces and Future Concerns," New Directions for Education, I (Spring, 1973).

Littell, W. M. "The Wechsler Intelligence Scale for Children: Review of a Decade of Research," Psychological Bulletin, LVII (1960), 132-56.

Livingston, Robert Tevlot. "The Theory of Organization and Management," Transactions of the ASME, May, 1953, p. 659.

McCammy, James L. "Analysis of the Process of Decision Making," Public Administration Review, VII (1947), 41.

Pate, R. H., Jr., and W. R. Nichols. "The Slosson Intelligence Test Used as a Screening Device: A Validity Study," American Psychological Association, Catalog of Abstracts, Experimental Publication System, 1970, No. 9, p. 10.

- Payne, James S. "Psychoeducational Diagnosis," Mental Retardation Courses at the University of Virginia. Charlottesville: University of Virginia, 1974.
- Roe, William H., and Thelbert L. Drake. The Principalship. New York: Macmillan, 1974.
- Sattler, Jerome M. Assessment of Children's Intelligence. Philadelphia: W. B. Saunders, 1974.
- Seashore, H. "Tenth Grade Tests as Predictors of Twelfth Grade Scholarship and College Entrance Status," Journal of Counseling Psychology, I (1954), 106-15.
- Silverstein, A. B. "Comparison of Two Item-Classification Schemes for the Stanford-Binet," Psychological Reports, XVII (1965), 964.
- Simon, Herbert A. Administrative Behavior. New York: Macmillan, 1950.
- \_\_\_\_\_. Administrative Behavior. 2d ed. New York: Macmillan, 1957.
- Sloan, W., and G. J. Bensberg. "An Exploratory Study of the Full Range Picture Vocabulary Test with Mental Defectives," American Journal of Mental Deficiency, LVIII (1954), 481-85.
- Slosson, Richard L. Slosson Intelligence Test for Children and Adults. New York: Slosson Educational Publications, 1963.
- Smith, L. M., and A. R. Fillmore. "The Ammons FRPV Test and the WISC for Remedial Reading Cases," Journal of Consulting Psychology, XVIII (1954), 332.
- Swanson, M. S., and A. Jacobson. "Evaluation of the Slosson Intelligence Test for Screening Children with Learning Disabilities," Journal of Learning Disabilities, III (1976), 318-20.
- Talmage, Harriet. Statistics as a Tool for Educational Practitioners. Berkeley, California: McCutchan, 1976.
- Terman, Lewis M., and Maud A. Merrill. Manual for the Stanford-Binet Intelligence Scale. 3d ed. Boston: Houghton Mifflin, 1973.
- Thorndike, Robert L., and Elizabeth Hagen. Measurement and Evaluation in Psychology and Education. 3d ed. New York: John Wiley and Sons, 1969.
- Torres, Scottie, ed. A Primer on Individualized Education Programs for Handicapped Children. Reston, Virginia: The Foundation for Exceptional Children, 1977.
- Tyler, L. E. The Psychology of Human Differences. 3d ed. New York: Appleton-Century-Crofts, 1965.

- Valett, Robert. Programming Learning Disabilities. Belmont, California: Fearon Publishers, 1969.
- Van Etten, Carlene, and Alen Van Etten. "The Measurement of Pupil Progress and Selecting Instructional Materials," Journal of Learning Disabilities, IX (October, 1976), 4.
- Vernon, P. E. Intelligence and Cultural Environment. London: Methuen, 1969.
- Wechsler, David. Manual for the Wechsler Intelligence Scale for Children-Revised. New York: Psychological Corporation, 1974.
- \_\_\_\_\_. The Measurement and Appraisal of Adult Intelligence. 4th ed. Baltimore: Williams and Wilkins, 1958.
- Wilson, John D., and Paul F. Spangler. "The Peabody Individual Achievement Test as a Clinical Tool," Journal of Learning Disabilities, VII (1974), 60-63.
- Wingo, G. Max. Philosophies of Education: An Introduction. Lexington, Massachusetts: D. C. Heath, 1974.

## APPENDIXES

**Appendix A**

**Procedures for the Identification, Evaluation,  
Confirmation, and Placement of Special  
Education Students**



# COUNTY OF ALBEMARLE



Department of Education  
ROOM 310, COUNTY OFFICE BUILDING  
CHARLOTTESVILLE, VIRGINIA 22901

OFFICE OF  
SUPERINTENDENT

## PROCEDURES FOR THE IDENTIFICATION, EVALUATION, CONFIRMATION, AND PLACEMENT OF SPECIAL EDUCATION STUDENTS

1. Referral of student by Classroom Teacher (on appropriate form) or by other appropriate persons and/or agencies. (Appropriate Elementary Supervisor should be contacted prior to referral by Classroom Teacher.)
2. Observation of student by LD Resource Teacher to obtain additional screening information regarding educational strengths and weaknesses and/or behavioral and/or emotional problems. Such screening information shall include existing routine vision, speech, language, and hearing exams, pre-school medical exams, and all group readiness, achievement, and intelligence tests.

A conference should be held with the parents to discuss the student's learning problems prior to a meeting of the School Building Screening Committee (see #3 below).

3. Referral to the School Building Screening Committee shall be made of the student whose screening records reveal:
  - a. Significant difference in academic performance (low) when compared with classmates;
  - b. Significant discrepancy between ability and achievement; or
  - c. Significant behavioral and/or physical problems.

The School Building Screening Committee shall consist of at least three. (3) of the following persons as appointed by the School Principal:

- a. Principal or designee
- b. Referring teacher
- c. Resource teacher (LD)
- d. Other teachers as appropriate
- e. Other building or county personnel as appropriate

Advancement to step 4 shall be made only if the screening committee decides that all appropriate resources and alternatives within the regular school

program have been exhausted in efforts to better meet the needs of the student and that the student is suspected of being handicapped.

4. LD Resource Teacher requests School Principal to obtain parent's permission on inclusive permission form entitled, "Permission for Evaluation" for educational, speech, language, and hearing, psychological, sociological, and health examinations of student.
5. Speech Clinician completes Speech, Language, and Hearing Report (if not already part of student's record).
6. LD Resource Teacher completes Educational Assessment (on proper form) of student.
7. LD Resource Teacher sends to Central Office (c/o Psychological Services) copy of Classroom Teacher Referral, Speech, Language, and Hearing Report, Educational Assessment, and Permission for Evaluation (signed by parent) and retains copy of same. COMPLETION OF THIS STEP DENOTES AN OFFICIAL SPECIAL EDUCATION REFERRAL.
8. The following procedures shall be completed simultaneously:
  - a. Student evaluated by or under the direct supervision of a Certified Psychologist
  - b. Parent of student is contacted by School Principal or designee for follow-up on completion and return (to school) of health examination (on proper form)
  - c. Parent of student is contacted by Visiting Teacher for completion of sociological. In addition, Visiting Teacher will facilitate (if necessary) completion of health evaluation by reminding parent that such needs to be done, which agencies can assist in having it done, and explaining the contents of the health form. This assistance by the Visiting Teacher does not relieve the parent and school from their responsibility of having the health form completed and returned to the school.
9. LD Resource Teacher is sent a copy of each of the evaluative components (psychological, sociological, and health examinations) as each is completed.
10. LD Resource Teacher completes and sends to Supervisor of Special Education form entitled, "Scheduling Eligibility Committee Agendas". A copy of this form is forwarded to School Principal.
11. Supervisor of Special Education checks completed form for Scheduling Eligibility Committee Agendas against Flow-Chart in Special Services Office. If form and chart agree with each other, Supervisor of Special Education schedules date for Eligibility Committee Meeting. Agenda of scheduled meeting is sent to School Principal one week prior to date of meeting.

12. Memo of results of Special Education Eligibility Committee Meeting is sent to School Principal on day following such meeting.
13. Special Education Eligibility Committee Report of Recommendations is sent to School Principal one week following committee meeting.
14. School Principal contacts parent to inform him/her of recommendations of Special Education Eligibility Committee and to invite him/her to participate in the development of the student's Individual Educational Program (IEP).
15. School Principal arranges meeting of the School Building IEP Committee, appointed by School Principal, which shall include, but is not limited to the following:

- \*a. School Principal or designee (chairperson)
- b. Teacher(s) (general and/or special)
- c. Parent and student (student only as appropriate)
- d. Other specialists as appropriate and designated by School Principal

\*If student's program is to be implemented in another school, participants of IEP Committee shall be selected as cooperatively arranged by the referring and receiving School Principals. It is recommended that the IEP Committee meet in the school where the student will be attending.

16. School Principal places student in special education program as determined by the agreed upon and signed IEP. A copy of the IEP is sent to Central Office (c/o Psychological Services).
17. Transportation arrangements are made by referring School Principal and Transportation Office for that student who will attend a different school as a result of new class placement.
18. All of the above evaluation components and related correspondence and information shall be filed in the student's Pupil's Confidential Record (green folder) and placed in the office of the School Principal or Guidance Counselor. A light green card shall be placed in the student's Pupil Cumulative Record (manila folder) to indicate the existence of the confidential record. A duplicate copy of this information shall be maintained in the Special Services Office.

## **Appendix B**

### **Procedures for the Reevaluation of Special Education Students**

# COUNTY OF ALBEMARLE

120



Department of Education  
ROOM 310, COUNTY OFFICE BUILDING  
CHARLOTTESVILLE, VIRGINIA 22901

OFFICE OF  
SUPERINTENDENT

## PROCEDURES FOR THE REEVALUATION OF SPECIAL EDUCATION STUDENTS

1. Special Education Teacher requests School Principal to obtain parent's permission form entitled, "Permission for Reevaluation" for educational, speech, language, and hearing, psychological, sociological, and health examinations of student. (If LD reevaluation, LD Resource Teacher makes this request of School Principal.)
2. Special Education Teacher completes Classroom Teacher Referral Form which is the reevaluation educational assessment. (If LD reevaluation, LD Resource Teacher completes the form;)
3. Special Education Teacher obtains Speech, Language, and Hearing Examination of student. (If LD reevaluation, LD Resource Teacher obtains examination.)
4. Special Education Teacher gives to LD Resource Teacher copy of Speech, Language, and Hearing Report, Classroom Teacher Referral Form, and Permission for Evaluation (signed by parent) and forwards to Central Office c/o Psychological Services. (If LD reevaluation, LD Resource Teacher obtains and forwards these items to Central Office c/o Psychological Services.)
5. The following procedures shall be completed simultaneously:
  - a. Student evaluated by or under the direct supervision of a Certified Psychologist
  - b. Parent of student is contacted by School Principal or designee for follow-up on completion and return (to school) of health examination (on proper form)
  - c. Parent of student is contacted by Visiting Teacher for completion of sociological. In addition, Visiting Teacher will facilitate (if necessary) completion of health evaluation by reminding parent that such needs to be done, which agencies can assist in having it done, and explaining the content of the health form. This assistance by the Visiting Teacher does not relieve the parent and school of responsibility for having the health form completed and returned to the school.
6. LD Resource Teacher is sent copy of each of the evaluative components (psychological, sociological and health examinations) as each is completed.

7. LD Resource Teacher completes and sends to Supervisor of Special Education form entitled, "Scheduling Eligibility Committee Agendas". Copies of this form are forwarded to the Special Education Teacher and the School Principal. LD Teacher sends evaluative components to appropriate Special Education Teacher. (If LD reevaluation, copy of form goes to the School Principal only.)
8. Supervisor of Special Education checks completed form for Scheduling Eligibility Committee Agendas against Flow-Chart in Special Services Office. If form and chart agree, Supervisor of Special Education schedules date for Eligibility Committee Meeting. Agenda of scheduled meeting is sent to School Principal and Special Education Teacher one week prior to date of meeting. (If LD reevaluation, agenda is sent to LD Resource Teacher rather than Special Education Teacher.)
9. Memorandum of results of Special Education Eligibility Committee Meeting is sent to School Principal on day following committee meeting.
10. Special Education Eligibility Committee Report of Recommendations is sent to School Principal one week following committee meeting.
11. School Principal contacts parents to inform them of recommendations of Special Education Eligibility Committee. If student is to remain in the special education program the School Principal invites the parents to participate in updating the Individual Educational Program (IEP). If student is to be returned to the regular program the School Principal informs the parents of such and receives parental permission on proper form entitled, "Parental Permission to Return to Regular Education From Special Education". Procedures and plan to return student to regular school program shall be cooperatively arranged during meeting of Eligibility Committee and follow-up meetings at the school.
12. School Principal arranges meeting of the School Building IEP Committee, appointed by School Principal which shall include, but is not limited to the following:
  - \*a. School Principal or designee (chairperson)
  - b. Teacher(s) (general and/or special)
  - c. Parent and student (student only as appropriate)
  - d. Other specialists as appropriate and designated by School Principal

\*If student's program is to be implemented in another school, participants of IEP Committee shall be selected as cooperatively arranged by the referring and receiving School Principals. It is recommended that the IEP Committee meet in the school where the student will be attending.

13. School Principal places student (or student remains in current program if no program change is indicated) in special education program as determined by the agreed upon, and signed IEP. School Principal forwards copy of IEP to Central Office.
14. Transportation arrangements are made by referring School Principal if student attends a different school as a result of new class placement.
15. All of the above related correspondence and information shall be filed in the student's Pupil Confidential Record (green folder) and placed in the office of the School Principal or Guidance Counselor. A duplicate copy of this information shall be maintained in the Special Services Office.

**Appendix C**  
**Collected Data Worksheet**



Time Code:  
A-Oct. - Dec.15  
B-Jan. - Feb.15

C-Feb.16 - Mar.31

**PROFILE SHEET**

**School:**

IO/MA

GE/MA

\*IQ not given takes higher one

## VITA

**Personal Data:**     Date of Birth: September 11, 1937  
                         Place of Birth: Harrisonburg, Virginia  
                         Marital Status: Married

**Education:**        Public Schools, Charlottesville, Virginia  
                         St. Andrews Presbyterian College, Laurinburg, North  
                         Carolina; English, Education, Psychology, B.A., 1963.  
                         University of Virginia, Charlottesville, Virginia;  
                         Counselor Education (Guidance, Counseling, Psychology),  
                         Special Education, Administration, M.ED., 1968.  
                         D.A.G.S. (Diploma of Advanced Graduate Study) same  
                         areas as M.ED., 1974.

**Professional  
Experience:**        Teacher, coach, athletic director, assistant director  
                         of music, Hargrave Military Academy; Chatham, Virginia,  
                         1963-1969.  
                         School psychologist, Central Office Staff of the Special  
                         Services Department, Albemarle County Public Schools;  
                         Charlottesville, Virginia, 1969-Present.

**Publications:**     Davis, Jr., Robert M., "Diagnostic Accountability,"  
                         International Journal of Instructional Media, New York:  
                         Baywood Publishing Company, Volume 5, Number 3, 1977-78,  
                         pp. 277-279.

**Honors and  
Awards:**            Music scholarship to outstanding musician, St. Andrews  
                         Presbyterian College.  
                         Phi Delta Kappa, University of Virginia.  
                         Kappa Delta Pi, University of Virginia.  
                         Doctoral Fellow, East Tennessee State University.